May 2019

SURFACE WATER – GROUNDWATER INTERACTIONS IN A PROGLACIAL ALPINE CATCHMENT: Applications of Heat Tracing, Modeling, and Remote Sensing Methods

Emily Alyssa Baker
Syracuse University

General Abstract

As climate change occurs, the availability and stability of water resources will be a global concern. The availability of reliable water resources will be of particular concern to regions that currently depend on meltwater from snowpack or glaciers during dry periods. One place whose water stability could be heavily impacted by the loss of meltwater resources is Peru. During the dry, arid months of May through September, the region to the west of the Peruvian Andes relies on stream water resources that originate in proglacial alpine catchments. The Cordillera Blanca, which is a mountain range in the Peruvian Andes, contains the highest density of tropical alpine glaciers worldwide, and the meltwater from these glaciers helps to sustain streamflow during the dry season. Unfortunately, tropical alpine glaciers are rapidly retreating as average annual air temperatures increase. In the Cordillera Blanca, glacial extent has decreased by more than 30% since 1930, and many of the glaciers have already passed a stage known as 'peak water', after which they continually contribute less and less water to streamflow. Recent research has indicated that although meltwater is a dominant contributor to streamflow during the dry season, groundwater within alpine aquifer systems may also be an important source of streamflow. Therefore, this research sought to investigate surface water – groundwater interactions in an alpine catchment of the Cordillera Blanca to gain a better understanding of the importance of groundwater in such regions.

This research focuses on the Quilcayhuanca Valley, which is a representative proglacial alpine catchment in the Cordillera Blanca, Peru. Glacial meltwater and groundwater contribute to streamflow within this catchment during the dry season. The Quilcayhuanca stream, along with streams that drain from adjacent catchments, flows into the Rio Santa which is a major stream in the region from which people withdraw water for a variety of uses. The high altitude wetlands in

the Quilcayhuanca Valley and similar catchments, known as pampas, may represent an important source of groundwater to streamflow. The valley aquifers consist of a mixture of landslide and talus slope deposits from the steep, adjacent bedrock cliffs and glaciofluvial deposits. The valley aquifers are confined from above by glaciolacustrine sediments that were deposited when proglacial lakes were present. In order to better understand the surface water – groundwater interactions in such a setting, we have combined energy balance modeling of stream heat fluxes, thermal infrared remote sensing of stream temperatures, and groundwater flow modeling to a section of the Quilcayhuanca stream that is downstream of direct glacial melt inputs to examine the influence of groundwater.

Energy balance modeling of stream temperatures, also known as heat tracing, can be used to estimate groundwater contribution to a stream. This method uses the fluxes of energy into and out of a stream to calculate stream temperature through time and space, and the difference between calculated and observed stream temperatures can be attributed to groundwater inflow. Meteorological and longitudinal stream temperature data were recorded for approximately a week in a portion of the Quilcay stream, and used as input for an energy balance model of the reach. Various model inputs were also varied in order to assess the sensitivity of the model to certain parameters and determine the extent to which uncertainty in certain model parameters affects estimates of groundwater influx. An input that was investigated in depth was the extent to which uncertainty in the daily diurnal streamflow signal affects groundwater inflow estimates, since the streamflow in proglacial streams varies diurnally due to glacial melt. Groundwater influx to the model reach was estimated at 42.1 L s^{-1} km^{-1}, and the uncertainty in diurnally fluctuating streamflow was determined to affect the estimated relative groundwater contribution to streamflow by approximately ±5%.

In order to improve the spatial resolution of the stream temperature data used to inform the energy balance modeling of stream temperatures, we explored the use of ground-based, time-lapse infrared remote sensing to measure stream temperatures along the same study reach in the Quilcayhuanca Valley. During two field seasons, a thermal infrared (TIR) camera was deployed on the steep valley cliffs, recording time-lapse images of stream temperature. Analysis of the infrared images revealed that measured infrared stream temperatures are highly sensitive to infrared temperatures from other objects in the environment that reflect off the stream surface, often leading to large discrepancies between in-situ and remote temperature measurements. We determined that even at nadir views, reflected temperatures can still affect measured TIR stream temperatures, and that previous analytical correction methods performed by the hydrology community have not accurately represented reflected infrared temperatures. While such analytical correction methods could be improved through more accurate measurements of reflected temperatures, an empirical correction approach can also be used to correct stream infrared temperatures. While this data was not ultimately used to refine the stream temperature energy balance model of the Quilcay stream, this investigation has helped improve our understanding of remote sensing of stream temperatures using ground-based, time-lapse thermal infrared imagery.

To complement the groundwater influx rates to the Quilcayhuanca stream estimated by stream temperature energy balance modeling, a groundwater flow model of the same pampa aquifer system was developed. Precipitation, stream stage, and groundwater level data that had been collected over various time periods were compiled to parameterize and calibrate the model. Numerous model simulations were explored to determine which model configuration best reproduced the annual hydraulic head patterns in the aquifer and the estimated dry season

groundwater flux to the reach of the Quilcay stream. The modeled groundwater flux estimates were then used to estimate the amount of groundwater entering the stream from all pampa regions of the catchment above the gauging station. Results indicate that about 7-53% of Quilcayhuanca streamflow is derived from groundwater depending on the month during the dry season. As the dry season progresses, the relative contribution of groundwater to the stream decreases as the aquifer becomes depleted. Travel time analysis also indicates that the residence time of water in the pampa aquifer system is relatively short, with >80% of water moving through the system in 6 months and the remaining water exiting after around 1-1.5 years. These results suggest that groundwater within these proglacial alpine catchments also contributes to streamflow, and that streamflow is vulnerable glacial loss, especially at the end of the dry season.

SURFACE WATER – GROUNDWATER INTERACTIONS IN A PROGLACIAL ALPINE CATCHMENT:

Applications of Heat Tracing, Modeling, and Remote Sensing Methods

by

Emily Alyssa Baker

B.A. Geology, Mount Holyoke College, 2015

Dissertation

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Earth Sciences

Syracuse University
May 2019

Acknowledgements

I would like to thank my advisor, Dr. Laura Lautz, for supporting me throughout my graduate career and providing me with the guidance I needed throughout the process. I would also like to thank Drs. Christa Kelleher, Jeffrey McKenzie, Laura Condon, and Zunli Lu for serving on my dissertation committee. I would also like to thank our research team with whom we work in Peru, including Bryan Mark, Robert Hellström, Lauren Somers, Robin Glas, Oliver Wigmore, Michel Baraer, and Pierrick Lamontagne-Hallé and Caroline Aubry-Wake. Without their long term effort and support the field work for this research would not have been possible. Additional thanks to Patricia Ames in Huaraz, Peru for her support between our field excursions. Lastly, I am incredibly grateful for the support of my family and friends, especially from my sister Jessica Baker and from Joseph Gonzalez, who have provided me with encouragement throughout the process.

I would also like to acknowledge the National Science Foundation through the Division of Earth Sciences (EAR-1316429), the Syracuse University Education Model Program on Water-Energy Research (National Science Foundation Grant DGE-1449617), the Syracuse University Water Fellowship, and the Central New York Association of Professional Geologists Student Research Grant for financial support of this work.

Table of Contents

List of Tables

Chapter 1

Chapter 2

Chapter 3

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Movies

Chapter 2

Chapter 1

The Importance of Incorporating Diurnally Fluctuating Stream Discharge in Stream Temperature Energy Balance Models

Published as:

Baker, E.A., L.K. Lautz, C. Kelleher, J.M. McKenzie. 2018. The importance of incorporating diurnally fluctuating stream discharge in stream temperature energy balance models. *Hydrological Processes*, 32, 2901-2914. https://doi.org/10.1002/hyp.13226

Abstract

Although stream temperature energy balance models are useful to predict temperature through time and space, a major unresolved question is whether fluctuations in stream discharge reduce model accuracy when not exactly represented. However, high-frequency (e.g., subdaily) discharge observations are often unavailable for such simulations, and therefore, diurnal streamflow fluctuations are not typically represented in energy balance models. These fluctuations are common due to evapotranspiration, snow pack or glacial melt, tidal influences within estuaries, and regulated river flows. In this work, we show when to account for diurnally fluctuating streamflow. To investigate how diurnal streamflow fluctuations affect predicted stream temperatures, we used a deterministic stream temperature model to simulate stream temperature along a reach in the Quilcayhuanca Valley, Peru, where discharge varies diurnally due to glacial melt. Diurnally fluctuating streamflow was varied alongside groundwater contributions via a series of computational experiments to assess how uncertainty in reach hydrology may impact simulated stream temperature. Results indicated that stream temperatures were more sensitive to the rate of groundwater inflow to the reach compared with the timing and amplitude of diurnal fluctuations in streamflow. Although incorporating observed diurnal fluctuations in discharge resulted in a small improvement in model RMSE, we also assessed other diurnal discharge signals and found that high amplitude signals were more influential on modelled stream temperatures when the discharge peaked at specific times. Results also showed that regardless of the diurnal discharge signal, the estimated groundwater flux to the reach only varied from 1.7% to 11.7% of the upstream discharge. However, diurnal discharge fluctuations likely have a stronger influence over longer reaches and in streams where the daily range in

discharge is larger, indicating that diurnal fluctuations in stream discharge should be considered in certain settings.

1 Introduction

Stream temperature can reflect various hydrologic processes including groundwater–surface water exchange, inflow of polluted waters, and mixing of source waters (e.g., Cassie, 2006; Dugdale, 2016; Gu, Montgomery, & Austin, 1998). Different water sources may exhibit different temperature signals that facilitate heat tracing of water flux rates. Temperature observations can be compared with simulated temperatures from a stream energy balance model to improve understanding of stream processes and their relative influence on simulated stream temperatures (e.g., Webb & Zhang, 1997; Westhoff et al., 2007). Because heat is conserved, changes in water temperature can be attributed to changes in heat fluxes to and from the stream or from the introduction of water from another source (e.g., Cassie, 2006; Somers et al., 2016; Webb & Zhang, 1997). Energy balance stream temperature models incorporate variables including air temperature, solar radiation, discharge rate, and morphologic characteristics such as channel shape to resolve stream temperature from calculated heat fluxes to and from the stream through space and/or time (e.g., Cassie, 2006; Glose, Lautz, & Baker, 2017; Loheide & Gorelick, 2006; Webb & Zhang, 1999; Westhoff et al., 2007). Examples of open source energy balance models include HeatSource (Boyd & Kasper, 2003), HFLUX (Glose et al., 2017), and many others (e.g., Dugdale, Hannah, & Malcolm, 2017; Kim & Chapra, 1997; King, Neilson, Overbeck, & Kane, 2016; Westhoff et al., 2007; Yearsley, 2009).

Diurnally varying discharge may be important to accurately represent in model simulations to accurately infer the relative importance of different processes on simulated stream temperatures. Although diurnal fluctuations in stream discharge may be important for accurately

determining stream temperatures using energy balance models, the diurnal discharge signal may not always be well known due to the typical low temporal resolution of field data. In ungauged streams, manual streamflow measurements may not reveal a diurnal signal, much less accurately capture the amplitude and time of peak discharge. Uncertainty in discharge measurements can be large enough to mask small diurnal signals because these types of measurements can have errors of up to 20%. Additionally, in reaches where diurnal fluctuations in stream discharge are expected, the daily signal may be difficult to predict, especially at downstream locations, if the cause of the fluctuations is far upstream (e.g., snowpack or glacial melt). Also, the cause of fluctuations will affect the timing of peak discharge, with melt-influenced discharge peaking late in the day depending on the distance downstream, and evapotranspiration-influenced discharge reaching a minimum in the afternoon. Causes of diurnal streamflow fluctuations include evapotranspiration (Deutscher et al., 2016; Wondzell, Gooseff, & McGlynn, 2010), estuary tidal cycles, snow/glacial melt (Greimel et al., 2016; Loheide II & Lundequist, 2009), agricultural influences (Younus, Hondzo, & Engel, 2000), dam operations, and wastewater effluent discharge (Greimel et al., 2016).

Despite their frequent occurrence in natural and impacted streams, there have been few assessments of how uncertainties in diurnally fluctuating discharge affect modelled stream temperatures. It is also unknown how such uncertainty compares to that generated by uncertainty in other model inputs. The effect of diurnal discharge fluctuations on energy balance model results is difficult to predict from first principles because as the amount of water in the stream changes, the width and depth of the stream also change, affecting the exchange of heat to and from the stream (e.g., Dugdale et al., 2017; Schmadel, Neilson, & Heavilin, 2015). Because many of the largest heat fluxes in energy balance models (King et al., 2016; Younus et al., 2000)

are ultimately scaled by surface area, it is important to know how stream width may change through time as discharge fluctuates, because the stream width will affect the magnitude of radiant heat fluxes across the air–water interface (King et al., 2016). In particular, the relative timing of peak stream widths and peak radiative heat fluxes may alter the extent to which changes in stream width impact energy balance model results. Channel geometry is also important because the wetted surface area to volume ratio scales with the water parcel's residence time within each model grid cell, and therefore the concurrent heat exchanges. Furthermore, as the discharge changes, the thermal mass is impacted, causing smaller temperature changes to result from the same heat fluxes as water volume increases. Additionally, changes in discharge affect streamflow velocity and thus residence time (Gu et al., 1998; Wondzell, Gooseff, & McGlynn, 2007), altering the heat exchange within a model grid cell, especially the amount of solar heating (Schmadel et al., 2015).

In this study, we examine the effects of diurnally fluctuating streamflow on calculated stream temperature by building on recent work. A previous study that considered only static discharge found substantial groundwater inflow to a 3,925-m reach of the Quilcay stream in Peru (Somers et al., 2016). Heat tracing results indicated ~29% of the stream discharge at the reach outlet came from groundwater discharge along the reach. The study also found differences in gross water exchanges and net groundwater inflow rates between different morphological sections of the reach. Dye tracing results indicated that gross stream–groundwater exchanges occurred within high-slope moraine portions of the reach, whereas net gains in streamflow from groundwater inflow occurred in low-slope meadows (Somers et al., 2016). However, the stream temperature energy balance model did not incorporate diurnal discharge fluctuations and instead assumed constant discharge through time. The 24-hr daily discharge pattern in this stream is

typical for regions exhibiting snow or glacial melt (Loheide II & Lundequist, 2009). Calculated stream temperatures and estimates of groundwater inflow may be affected by modelling streamflow as constant through time, but this remains unexplored.

The current study sought to determine if excluding diurnal fluctuations in streamflow affects calculated stream temperatures within a portion of the Quilcay stream that traverses a meadow. Stream temperatures were calculated using the HFLUX Stream Temperatures Solver, a deterministic energy balance model for calculating stream heat fluxes. Our research goal is to assess the uncertainty in predicted stream temperatures and inferred groundwater discharge rates introduced by failure to account for diurnal fluctuations in streamflow. Through the use of model simulations across a range of input values chosen using both random (Monte Carlo) and nonrandom sampling schemes, we explore how variability in diurnally fluctuating streamflow driven by different hydrologic processes may impact stream temperature simulations in stream reaches.

2 Study Area

The field site for this study is in the Quilcayhuanca Valley of the Cordillera Blanca, Peru, which features some of the most heavily glacierized mountains in the tropics (Suarez, Chevallier, Pouyaud, & Lopez, 2008). Quilcayhuanca Valley is typical of the low-slope meadow systems within long, deep, hanging valleys of the Cordillera Blanca (Gordon et al., 2015). Precipitation is seasonal, though annual air temperature experiences little variability, with daily air temperatures fluctuating more than average annual air temperatures (Kaser, Ames, & Zamora, 1990). This causes glacial ablation to occur throughout the year, rather than seasonally, with greater glacial ablation occurring during the wet season than the dry season (Kaser et al., 1990; Mark & Seltzer, 2003; Suarez et al., 2008). Melt water from these glaciers causes proglacial streams in the region

to exhibit diurnal fluctuations in streamflow. The Quilcayhuanca study reach is located over 6.7 km downstream of the glacier, resulting in a delay in the timing of the peak of these diurnal fluctuations. The Quilcay stream is a proglacial, alpine stream that exhibits diurnal fluctuations in stream discharge that peak in the evening (~7:00 p.m.) within the study reach and fluctuate by ~10% of the mean discharge daily.

The input data for the HFLUX stream temperature model (Glose et al., 2017) was collected in a ~1.2-km reach of the Quilcay stream (Figure 1a), which flows through a low-gradient portion of the valley floor, surrounded by steep valley walls. Most of the reach has no overhanging vegetation, except for a few tens of meters at the upper end of the reach where small trees overhang the stream. The weather during the dry season is consistent from day to day, exhibiting strong diurnal patterns in incoming shortwave radiation (Figure 1b), air temperature (Figure 1c), relative humidity (Figure 1d), and wind speed (Figure 1e). Incoming shortwave radiation is high, with maximum values exceeding 900 W m^{-2}, and cloud cover is typically low. The average groundwater temperature in five wells within the reach was 10.6°C during the model period, with individual wells having average temperatures ranging from 9.0°C to 11.3°C. The streambed is armored by cobbles that do not appear to be infilled with fines. Stream channel width ranges from ~3 to ~14 m across in the widest, braided section, with a mean width of 6.1 m. The mean channel depth is 0.3 m, and the mean discharge during the study period was 0.79 m^3 s^{-1}. The observed range in bank angles was from 143° to 177° (width relative to depth), with a median angle of 168° (Figure 2).

3 Methodology

3.1 Field Methods

To estimate the influence of diurnally fluctuating streamflow on simulated stream temperatures, the HFLUX Stream Temperature Solver was used. The required input data for the model were collected in the Quilcayhuanca Valley (Figure 1). Stream temperatures in a ~1.2-km reach of the Quilcay stream (9.4656°S, 77.3792°W, ~3,930 m above sea level) were recorded from July 20 to July 26, 2015 using in-stream Thermochron iButtons (Model DS1922L). The iButton sensors were cross calibrated in a water bath. We deployed 40 iButton temperature sensors within the Quilcay stream at ~25-m intervals, which recorded data every 5 min with a resolution of 0.0625°C (manufacturer accuracy of ±0.5°C). The sensors were installed on stakes vertically oriented in the stream, and reflective tape was placed around them in order to minimize exposure to direct solar radiation. We also installed a series of three sensors vertically within the water column to obtain temperature profiles at three locations within the reach to confirm that the stream was well mixed. Temperature sensors were also installed in the streambed at depths of 15 and 25 cm below the water-streambed interface to measure streambed temperatures. The average temperatures through time recorded by the sensors at the 15-cm depth were used as the streambed temperatures in the model. Stream depths at three points spaced across the stream were measured at each of the sensor locations. Stream widths along the reach were measured using 10-cm resolution aerial imagery of the study site (Wigmore & Mark, 2017). The bed material mainly consisted of large cobbles, and this sediment size was used to assign the value of sediment thermal conductivity (see Glose et al., 2017).

Forcing data for the stream temperature model were collected via a Vantage Pro2 weather station in approximately the center of the study area near the stream. Weather data, including air

temperature (°C), relative humidity (%), wind speed (m s^{-1}), and incoming solar radiation (W m^{-2}), were recorded at 10-min intervals from July 20 to July 26 (Figure 1b–1e). Cloudiness was estimated by comparing incoming solar radiation at a given time of day to the solar radiation at that time on a day of full sun (July 25; Figure 1b). The percentage of stream shading at multiple locations along the reach was estimated by measuring stream width and the portion of that width covered by vegetation using 10-cm resolution aerial imagery (Wigmore & Mark, 2017). These shading percentages were used to estimate the view to sky along the reach as 1-shading. Groundwater temperatures were recorded hourly via Schlumberger Mini-Diver pressure loggers. River stage data were downloaded from a pressure transducer recording water height in a flume within the study reach (Figure 1g). The stage data were corrected by subtracting the barometric pressure and then converting the stage to discharge using a rating curve produced for the Casa de Agua flume (Q(L s^{-1}) = 6145.7h2–411.48h + 330.1; R^2 = 0.9, n = 7). Three manual discharge measurements were also taken during the 5-day observation period using a Marsh McBirney Flo-Mate 2000 Flow Meter at approximately the midpoint of the study area; two measurements were taken during mornings, and one during an evening (Figure 1g). The error associated with these measurements is ±20% of the measured discharge.

3.2 Modelling Methods

The HFLUX Stream Temperature Solver (http://hydrology. syr.edu/hflux/), a 1D longitudinal stream energy balance model programmed in MATLAB (Glose et al., 2017), was used to model stream temperature for a reach of the Quilcay stream. The required inputs for this model are listed in Glose et al. (2017) and were measured during July 2015 (see Section 3.1). HFLUX was chosen because it can easily be adapted to different systems, modelling objectives, and repeated model runs by altering the MATLAB code. HFLUX is a Eulerian deterministic

stream temperature model that calculates heat fluxes at specified grid cell locations using a finite difference method and then uses those fluxes combined with the groundwater inflow at each cell to determine the stream temperature through both time and space (see Glose et al., 2017, for a more detailed model description). HFLUX uses mass and energy balance equations common to many different energy balance models (e.g., Boyd & Kasper, 2003; Kim & Chapra, 1997; Webb & Zhang, 1997; Westhoff et al., 2007). Detailed descriptions about the specific equations used and methods of calculation for each heat flux, along with additional model details and required inputs, can be found in Glose et al. (2017).

Heat fluxes calculated within HFLUX include net shortwave radiation, sensible heat flux, latent heat flux, streambed conduction, and net longwave radiation (taken as the sum of the atmospheric longwave radiation, back radiation off the stream, and landcover radiation; Figure 1h). For the model simulations in this study, the Crank–Nicolson method was used to solve the finite difference equations, the mass transfer method (HFLUX default empirical constants) was used to calculate the latent heat flux, Bowen's ratio was used to calculate sensible heat from the latent heat flux, and measured incoming shortwave radiation was corrected for shading and reflection to calculate net shortwave radiation (e.g., Boyd & Kasper, 2003; Glose et al., 2017; Kim & Chapra, 1997; Webb & Zhang, 1999; Westhoff et al., 2007). The HFLUX code allows stream discharge to vary through time by relating the velocity–discharge equation and the Manning equation (Gu et al., 1998), such that as the discharge (Q) increases, the channel shape (θ), slope (s), and roughness (n) are held constant while the width (W) and depth (D) increase according to their determined relationship

to discharge (Figure 2):

$$D = \left(\frac{2 \times n \times Q}{3.57 \times cos\theta^{2/3} \times tan\theta^{5/3} \times s^{1/2}}\right)^{3/8} \quad [1]$$

$$W = 2 \times tan\theta \times \left(\frac{2 \times n \times Q}{3.57 \times cos\theta^{2/3} \times tan\theta^{5/3} \times s^{1/2}}\right)^{3/8} \quad [2]$$

Using this approach, we varied stream width with discharge, consistent with studies that have found that stream temperatures are sensitive to the channel geometry, especially at low flows, due to the scaling of the heat fluxes by the surface area to volume ratio of model grid cells (e.g., Schmadel et al., 2015). The HFLUX code assumes a triangular channel shape (Glose et al., 2017), and prior work by the authors in the development of HFLUX indicates that the results for this system are not sensitive to a rectangular versus triangular channel shape. Modelled stream temperatures were calculated every minute at 1.0-m increments along the reach.

3.3 HFLUX Simulation Scenarios

Monte Carlo analyses of HFLUX were run in MATLAB with gradually increasing complexity to assess the uncertainty associated with model inputs (Wagener & Kollat, 2007), including diurnal discharge patterns, channel geometry, and the rate and temperature of groundwater inflow. These analyses help us understand how uncertainties in these inputs may affect simulated spatio-temporal stream temperatures. Patterns in diurnal discharge fluctuations are described by their amplitude, timing (e.g., of maximum or minimum flow), and average discharge rate. These simulations enabled us to assess the ranges of uncertainty in these inputs that still yielded acceptable modelled temperature results. We also determine how uncertainty in the diurnal discharge pattern interacts with uncertainties in other model parameters to affect model output (Pianosi, Sarrazin, & Wagener, 2015; Wagener & Kollat, 2007). The parameters varied include timing of peak discharge, amplitude of diurnal discharge fluctuation, mean discharge through time, channel shape, and the rate of groundwater inflow (Figure 3). Ranges for these inputs are listed in Table 1.

We carried out five different sets of Monte Carlo analyses: (a) varying diurnal discharge timing and amplitude; (b) varying diurnal discharge timing, amplitude, and mean flow; (c) varying diurnal discharge timing and amplitude along with the channel shape (bank angle, Figure 2); (d) varying diurnal discharge timing and amplitude along with the groundwater flux; and (e) varying the groundwater flux and the groundwater temperature. These sets of analyses were run to gradually increase model complexity. In this way, we tested the effects of the different inputs systematically to assess and limit interactions. For each analysis, we randomly sampled the input space across ranges listed in Table 1 to create input sets of 500 simulations. Simulated discharges were modelled using a sine function, with discharge smoothly fluctuating over a 24-hr period (Wondzell et al., 2007) and the amplitude and timing of peak discharge randomly chosen for each simulation (Figure 3). Minimum discharge was assumed to occur 12 hr after peak discharge. Base case models were also run using the measured diurnal discharge smoothed using a moving average to remove the noise and using constant discharge through time (the mean of the observed discharge; Figure 3). Base case models did not include groundwater inflow. The upstream temperature boundary condition was not impacted by the temperature of melt at the glacier toe because the source of the glacial melt water is far enough upstream to equilibrate with atmospheric heat fluxes prior to entering the study reach. The input data were not adjusted to improve the model fit and reduce the error. Input data for the base case models are accessible from the CUAHSI HydroShare data repository (Baker, Lautz, Kelleher, & McKenzie, 2018).

Model errors were computed using the five most downstream observation points as they are furthest from the upstream boundary and therefore least impacted by initial boundary conditions. Model error was quantified using the root mean square error (RMSE) and the

absolute differences between observed and simulated temperatures. The temperature differences at the (a) maximum, (b) minimum, (c) mid-increasing (midmorning), and (d) mid-decreasing (evening) stream temperatures were also used to assess error.

We also performed a set of 1,248 non-randomly sampled model runs to determine the optimal groundwater inflow rate across the range of diurnal discharge signals. The optimal groundwater inflow rate under each diurnal discharge timing and amplitude scenario was the rate that minimized the downstream RMSE. Sixty different diurnal discharge timing and amplitude combinations were used. For each of these scenarios, the model was run with different groundwater flux rates using the observed average groundwater temperature of 10.6°C. Sixteen groundwater flux rates (ranging from ±20% of the stream discharge) were used for each diurnal discharge scenario. This analysis was then repeated using groundwater temperatures of 4°C and 13°C.

4 Results

4.1 Base Case Model Results

Improvements in model accuracy by incorporating temporally varying discharge were greatest when stream temperatures were decreasing, especially on the evenings of the 24th to 25th and the 25th to 26th (Figure 4), suggesting the observed discharge hydrograph improves the timing, rather than the amplitude, of temperature changes. The stream temperatures simulated by HFLUX for average (time-constant) discharge yielded an overall RMSE of 0.27°C, with a range from 0.04°C to 0.64°C (Figure 4; Table 2). When the model was run using the observed hydrograph, which exhibited diurnal discharge fluctuations, the overall RMSE was 0.26°C, with values from 0.04°C to 0.58°C. The RMSE through time is similar for both the observed and

constant discharge scenarios, with the smallest RMSE typically occurring between 9:00 a.m. and 3:00 p.m. (Figure 4c). The mean temperature difference through time is also similar for the two scenarios, with stream temperatures typically underestimated at night by up to 0.64°C when discharge is constant and up to 0.57°C when discharge varies as observed (Figure 4d).

4.2 Discharge Timing and Amplitude Monte Carlo Simulations

Uncertainties in simulated stream temperature were largest for Monte Carlo experiments varying the rate of groundwater inflow as opposed to experiments that varied diurnal discharge timing, amplitude, average discharge, or channel geometry (Figure 5). Plots of the average difference in stream temperature through time were used to compare modelled temperatures to observed temperatures. Uncertainties in the timing and amplitude of diurnal discharge fluctuations resulted in simulated temperatures of at most 1.9°C warmer and up to 1.8°C colder than observed temperatures (Figure 5a). The largest errors tended to occur in the night to early morning (~1:00 a.m.–6:00 a.m.) and during the later morning (9:00 a.m.–12:00 p.m.), and temperatures were more frequently under-predicted than over-predicted (Figure 5a). When we additionally considered uncertainty in channel shape (alongside diurnal discharge timing and amplitude), the difference between observed and simulated stream temperatures were overestimated up to 3.4°C, or underestimated up to 2.5°C (Figure 5b). Adding uncertainty in the rate of groundwater inflow (alongside diurnal discharge timing and amplitude) yielded the largest model temperature deviations, with absolute differences of up to 7.4°C warmer than measured temperatures (Figure 5c). Adding in uncertainty in the mean discharge rate resulted in predicted temperatures of up to 3.1°C too warm, but typically too cool, with temperatures underestimated by up to 2.4°C (Figure 5d). Errors across all uncertainty scenarios were often similarly distributed in time, but larger in magnitude, than the two base case scenarios (Figure 5).

To test how timing of minimum discharge and amplitude of the diurnal discharge fluctuations affected simulated stream temperatures across the day, we calculated the average temperature differences at the maximum, minimum, mid-increasing (e.g., midmorning), and mid-decreasing (e.g., early evening) stream temperatures for the Monte Carlo model outputs (Figure 6). Peak stream temperatures (Figure 6a) and early evening stream temperatures (Figure 6d) had the smallest prediction error, with errors ranging from −0.01°C to 0.09°C and from −0.45°C to −0.18°C, respectively. Minimum stream temperatures (Figure 6b) and midmorning stream temperatures (Figure 6c) had the largest prediction error, with error ranging from −1.2°C to −0.07°C and from −0.11°C to 1.0°C, respectively. The minimum and early evening stream temperatures were always underestimated. The errors were largest and had the most variation when the diurnal discharge signal had a large amplitude (>30%), with the timing of the minimum diurnal discharge affecting the observed stream temperatures more at higher amplitudes, indicating an interaction between the timing and amplitude of the diurnal discharge signal.

The set of Monte Carlo simulations where the diurnal discharge timing and amplitude, along with the groundwater inflow rate, were varied shows that while there is an interaction between the diurnal discharge pattern and the rate of groundwater inflow, the rate of groundwater inflow dominates the resulting modelled temperatures (Figure 7). Models that best predict peak stream temperatures have slightly positive rates of groundwater inflow. The errors at peak temperatures are greatest when the amplitude is high and peaks in the early morning, especially under large discharge losses (Figure 7a,b). At these peak stream temperatures, the calculated temperatures become too cold as the groundwater influx increases and too warm as the stream loses more water to the subsurface (Figure 7a,b). Models of losing streams result in peak stream temperatures that are up to 5.5°C warmer than measured stream temperatures

(Figure 7a,b). The temperature differences at minimum stream temperatures have higher errors when the amplitude is high and peaks in the evening between 5:00 p.m.–10:00 p.m. (Figure 7c,d). Minimum temperatures are most accurate when the groundwater inflow rate is greater than about ±5% of the upstream discharge, with up to a −1.2°C temperature difference when the groundwater inflow is minimal (Figure 7c-d). Minimum stream temperatures are similarly sensitive to both groundwater gains and stream water losses along the reach (Figure 7c,d). When stream temperatures are increasing, the temperature differences are highest when the amplitude is high and peaks in the evening to night-time when the stream is strongly losing water to the subsurface (Figure 7e-f). When stream temperatures are decreasing, the temperature differences are highest when the discharge peaks in the morning at high amplitudes (Figure 7g,h). The temperature differences at these four different stream temperatures demonstrate that stream temperatures can either be overestimated or be underestimated, depending on the change in discharge and the diurnal discharge signal. When model error is described using RMSE, the error is greatest when the stream is losing water to the subsurface, especially when this loss occurs under high amplitude discharges that peak after midnight (Figure 7i,j).

4.3 Monte Carlo Analysis of Groundwater Temperature and Inflow Rate

The Monte Carlo analysis varying groundwater temperature and flux shows an interaction between these two inputs (Figure 8). The optimal combination of groundwater flux and groundwater temperature was determined for both the constant and diurnal discharge scenarios. A 51.8 L s^{-1} influx of 12.0°C groundwater was optimal for the constant discharge scenario (RMSE = 0.19°C), whereas a 50.1 L s^{-1} influx of 11.3°C groundwater was optimal when discharge fluctuated diurnally (RMSE = 0.18°C). When groundwater fluxes mimicked a losing reach, model error increased regardless of the groundwater temperature; when gaining,

both the temperature and rate of groundwater inflow increased model errors. Overall, the observed discharge scenario (Figure 8b) produced estimates of groundwater temperature that are slightly improved compared with those estimated from the constant discharge scenario (Figure 8a).

4.4 Non-randomly Sampled Stream Discharge and Groundwater Flux Analysis

For the non-randomly sampled analysis, model output was calculated across a range of diurnal discharge patterns for multiple rates of groundwater flux, using the observed average groundwater temperature of 10.6°C (Figure 9). This analysis indicated that regardless of the diurnal discharge timing and amplitude, the optimized groundwater flux to the stream reach was within ±5% of the flux predicted for the observed discharge pattern (6.7 ± 5.0% of the initial upstream discharge; Figure 9). The RMSE indicated that a groundwater inflow of 6.7% was the most likely (~32%) optimized value over the range of input combinations. When the model was run with no diurnal signal, a groundwater inflow of 6.7% was still optimal within the reach. About 42% of the discharge timing and amplitude combinations predict slightly smaller groundwater inflows ranging from 1.7% to 5.0%. These smaller inflow rates occur as the amplitude of the diurnal signal increases and when the discharge peaks at night between 9:00 p.m. and 7:00 a.m. Meanwhile, about 27% of the diurnal signals produced slightly larger estimates of groundwater influx (8.3–11.7%). These larger rates occur as the amplitude increases and peaks during the day between ~9:00 a.m. and 7:00 p.m. Both the observed diurnal stream discharge (~10% fluctuation peaking around 7:00 p.m.) and the constant discharge scenario predicted the same optimal groundwater contribution for the study reach (Figure 9). This analysis was also performed using groundwater temperatures that were colder (4°C) and warmer (13°C) than the stream temperatures (Figure S1). Although these simulations predicted different water

fluxes to or from the stream, the variation in predicted fluxes due to the range of diurnal streamflow signals was similar.

5 Discussion

5.1 Do Diurnal Fluctuations in Stream Discharge Impact Modelled Quilcay Stream Temperatures?

The model output from the constant discharge scenario and observed diurnal discharge scenario indicate that incorporating the diurnal discharge fluctuations into the energy balance model of this reach does not significantly alter the calculated stream temperatures, with both scenarios having similar errors through time (Figure 4). It does, however, yield marginal improvements in model error, improving RMSE by 0.01°C (Figure 4c; Table 2). Through time, the RMSE and average temperature differences between the modelled and measured stream temperatures particularly improved during decreasing stream temperatures when the observed diurnal streamflow was used (Figure 4c,d; Table 2). However, these changes are also very small, with the greatest improvement of ~0.1°C. We hypothesize incorporating diurnal discharge fluctuations may have a greater impact on simulated stream temperatures over longer model reaches, where the distance from the upstream model boundary is greater; these fluctuations may also have a larger impact in larger or regulated rivers. Following Heavilin and Neilson (2012a, 2012b), we found that the temperature boundary condition is ~90% of the total temperature at 1.2 km during the day and ~98% during the night. Similarly, Schmadel et al. (2015) found that the upstream boundary could have an effect on calculated stream temperatures at distances of over 25 km (Schmadel et al., 2015). However, the model results at the downstream boundary are not solely a function of the upstream boundary condition during the day, as they do respond to changes in the model inputs.

5.2 Are Stream Temperatures More Sensitive to Diurnal Streamflow Patterns or Groundwater Flux Rates?

The groundwater inflow rate to the reach is more influential on simulated stream temperatures than the timing and amplitude of diurnal streamflow fluctuations, according to the temperature differences through time (Figure 5a and 5c). Although the variability produced by the groundwater flux rate dominates over the smaller variability produced by the diurnal stream discharge signal, the timing and amplitude of the discharge still interact to produce differences in the calculated stream temperatures (Figures 6 and 7). At peak stream temperatures in losing streams, high diurnal discharge amplitudes can produce varying differences between the modelled and observed temperatures, depending on the timing of peak discharge. If the discharge peaks in the afternoon to evening (10:00 a.m.–10:00 p.m.), the error is lower, whereas if the discharge peaks in the early morning (12:00 a.m.–6:00 a.m.), the error is largest (Figure 7a,b). At minimum stream temperatures, this pattern is reversed, with high amplitude signals producing the smallest temperatures differences when the discharge peaks in the early morning and producing the largest temperature differences when it peaks in the evening, especially when the groundwater flux is minimal (Figure 7c,d). When RMSE is used, the greatest errors occur for losing streams when the amplitude is high and peaks after midnight (Figure 7i,j). When groundwater flux is excluded, peak stream temperatures are typically slightly overestimated (up to 0.09°C) but can be underestimated by 0.01°C when the amplitude of the diurnal discharge signal is large and peaks around noon (Figure 6a). This occurs because the stream width reaches a minimum when the stream should be warming, reducing the heat flux to the stream. Evapotranspiration could cause minimal discharge midday. However, it is unlikely that a stream would experience evaporative losses of over 30%, as such a phenomenon may only occur in

extremely shallow or ephemeral streams. Peak temperatures are probably typically slightly overestimated because groundwater inflow is not included in these model runs (Figure 6a).

Although minimum stream temperatures are always underestimated when groundwater flux is not included, the diurnal discharge scenario can increase the temperature difference between simulated and observed temperatures. The consistent underestimation of minimum stream temperatures may indicate that heat fluxes during night-time hours are not completely described by the model or that groundwater inflow needs to be included in the model to accurately simulate stream temperatures. Diurnal changes in streamflow amplify these minimum stream temperature errors when discharge is at a minimum in the early morning (3:00 a.m.–9:00 a.m.), representing a stream with minimal water volume during times of low heat flux and therefore experiencing more cooling. This scenario may also be most applicable to streams impacted by upstream melt water (Figure 1g). Therefore, if meltwater causes a 20% or greater daily change in the stream discharge, modelled minimum stream temperatures may be underestimated by >1°C if the diurnal discharge signal is excluded (Figure 6b). In a longer reach, these temperature differences could become even larger, making the incorporation of diurnal streamflow fluctuations more important over longer reach lengths.

Varying the rate of groundwater inflow to the reach, in addition to the diurnal discharge signal, added greater uncertainty to the modelled stream temperatures. Regardless of the error metric used, simulations of losing streams overestimated stream temperatures, except at minimum stream temperatures (Figure 7). In contrast, simulations of gaining streams yielded small errors at peak stream temperatures (Figure 7a) and minimized errors at cold stream temperatures (Figure 7 b). This result is in part due to groundwater temperatures (10.6°C) that are more similar to the daily average peak (12.2°C to 13.0°C) than daily minimum (4.3°C to

6.4°C) stream temperatures. Additionally, a losing stream will have increasingly low stream volumes with distance. At peak daytime stream temperatures, low stream volumes allow for greater solar heating during the day and so the heat fluxes from shortwave radiation will produce greater increases in stream temperature (Figure 7a). Meanwhile, in gaining streams, larger stream volumes minimize changes in temperature from solar heating (Figure 7a). Furthermore, in the Quilcay stream, groundwater temperatures (10.6°C) are closer to peak stream temperatures and so a gaining stream will experience minimal cooling from inflowing groundwater during the day, but warming from groundwater during the night.

These interactions between the diurnal discharge signal and the change in streamflow along the reach are due to the relative importance of advective and nonadvective heat fluxes during different scenarios. At minimum stream temperatures, the temperature of losing reaches is driven by nonadvective heat exchange and the temperature of gaining reaches is again driven by advective heat exchange (Figure 7c,d). At night, the small widths of losing streams minimize the heat fluxes sufficiently that the stream does not cool, despite a lower water volume. Conversely, during the night in gaining streams, the increase in groundwater flux heats the stream due to advective heat exchange with the relatively warm groundwater. The models with the smallest RMSE values are therefore a balance of the advective and nonadvective heat fluxes that result in accurate stream temperatures both at warm and cold stream temperatures.

5.3 Do Diurnal Streamflow Patterns Impact Groundwater Inflow Rates Estimated from Energy Balance Models?

Because stream temperatures are more sensitive to the groundwater inflow rate as compared with the timing and amplitude of the diurnal discharge fluctuation (Figure 5c), it is not surprising that optimization from the non-randomly sampled analysis estimated similar

groundwater inflow rates (1.7–11.7% change in discharge) for many of the timing and amplitude combinations (Figure 9). The observed amplitude of the Quilcay stream is about 10% of the mean discharge through time and peaks between around 6:00 p.m.–7:00 p.m. When the nonrandomly sampled simulations were run to determine the optimal groundwater flux along the reach, the optimal groundwater contribution is 50.3 L s^{-1} for both the constant and observed stream discharge scenarios. Therefore, in the Quilcay stream, incorporation of the diurnal discharge fluctuations within the HFLUX energy balance model (instead of assuming constant discharge through time) does not produce a difference in the estimated inflow of groundwater into the modelled reach, validating previous estimates of groundwater inflow made in this valley that did not incorporate the diurnal discharge signal (Somers et al., 2016). However, if the stream had a different diurnal discharge signal, its incorporation may have mattered. The largest difference between models with constant discharge and diurnal discharge within the explored space occurs when the amplitude of the diurnal discharge is high (40–50%) and peaks in the night (10:00 p.m.–4:00 a.m.) or afternoon (12:00 p.m.–6:00 p.m.). Under these diurnal signals, the optimal rate of groundwater inflow is as much as ±5% different than the flux predicted for the observed discharge (Figure 9). When simulations are run using other groundwater temperatures, the range of diurnal discharge signal still only produces uncertainty in the groundwater flux rate of up to ±5% of the upstream discharge (Figure S1). Therefore, incorporating the correct diurnal stream discharge makes minor differences in the rate of groundwater inflow predicted by the energy balance model. Thus, in streams whose discharge signal may be close to those yielding differences in predicted groundwater fluxes, an attempt should be made to measure and constrain the diurnal discharge fluctuations.

Incorporation of the diurnal discharge signal is likely to be more important in correctly estimating the groundwater contribution if the true amplitude of the diurnal signal is much higher than that observed in the Quilcay stream, for example, in streams where the discharge is regulated by dams and can fluctuate by 60% or more daily (e.g., Smokorowski, 2010). Another situation where the diurnal fluctuations may be more important is in a stream where the daily difference between the maximum and minimum stream temperatures is smaller and the groundwater temperature is within the range of daily stream temperatures. Under such conditions, the rate of groundwater inflow will have a smaller impact on modelled stream temperatures, potentially making incorporation of the diurnal discharge fluctuation more important for accurately estimating the groundwater contribution to the reach. Furthermore, over longer reach lengths than in this model (>1.2 km), the effects of diurnal variations in discharge on stream temperature may be greater further downstream, likely having a larger effect on the predicted groundwater contribution. This idea is supported by another study that examined the influence of channel geometry on stream temperature and found that it had a greater effect in a longer (25 km) reach due to the control of the upstream boundary conditions on modelled stream temperatures (Schmadel et al., 2015). Therefore, although incorporating the diurnal discharge signal may not be necessary to predict groundwater inflow at the Quilcayhuanca field site to within a few percent, it may be a more important consideration for streams in other locations with different characteristics.

The previous study in Quilcayhuanca Valley that used HFLUX to estimate the groundwater contribution to a reach of the stream analyzed how the groundwater temperature and inflow rate interact to produce estimates of groundwater inflow (Somers et al., 2016). The study used constant discharge through time within the model. The results from the current study

show a similar relationship between groundwater temperature and inflow rate, with the highest model errors occurring when the groundwater contribution is large and the groundwater temperature is warm (>13°C) or cold (<5°C; Figure 8). However, the optimized groundwater inflows predicted in this study (43.6 L s^{-1} km^{-1} for constant discharge, 42.1 L s^{-1} km^{-1} for observed discharge, Figure 8a,b) are lower than that predicted in Somers et al. (2016; 58.6 L s^{-1} km^{-1} inflow of 8.8°C). Meanwhile, the optimized groundwater temperature in this study (11.3°C) is slightly warmer than the observed 10.6°C average, whereas the Somers et al. (2016) optimized groundwater temperature of 8.8°C is lower than the measured 9.4°C average groundwater temperatures during their study. Because the average temperatures of the five wells ranged from 9°C to 11.3°C, the optimized groundwater temperature for the diurnal discharge scenario falls within the temperature range observed in this study. Additionally, the difference between the optimized and observed groundwater temperatures could be due to the uncertainty in the temperature of the groundwater that is actually entering the stream, as its temperature can change as it travels to the stream and through the stream bed (Kurylyk, Moore, & MacQuarrie, 2015). Meanwhile, differences between the optimized groundwater flux and temperature estimates for our study and the Somers et al. (2016) study could be due to the almost twice as long study reach in the Somers et al. (2016) study, which also flowed through a moraine deposit and a further upstream meadow. Furthermore, differences in groundwater temperatures between the two studies could be due to the different times that the studies were conducted or to the higher elevation of the upstream half of the Somers et al. (2016) study reach.

6 Conclusion

The outputs from the Monte Carlo and non-randomly sampled HFLUX model simulations show that for this reach of the Quilcay stream, incorporating the observed diurnal

discharge signal as opposed to assuming constant discharge through time does not affect the estimated rate of groundwater contribution to the stream. However, if the diurnal signal in the Quilcay stream had been larger (e.g., 40% daily variation peaking at 4:00 a.m.), the estimated groundwater inflow rates could have differed if constant discharge was assumed (3.3% vs. 6.7% of the upstream discharge). Therefore, slight differences in the estimated groundwater contribution determined from energy balance models can occur depending on the true diurnal discharge signal and whether it is approximated. Although the incorporation of diurnal discharge fluctuations did not significantly alter the estimated inflow of groundwater to the studied reach of the Quilcay stream, it may do so in reaches with different characteristics, such as those with stronger diurnal discharge signals, or over longer model reaches. Additional uncertainty in the groundwater temperature could lead to larger differences between the groundwater contribution to the reach estimated from constant verses diurnal discharge models, especially depending on the metric used to assess model performance and behavior. Although leaving out the diurnal discharge fluctuations within a stream may not greatly affect estimates of the groundwater inflow rate, it can affect predicted stream temperatures (Figure 10), especially further from the upstream boundary. This could be important to those using energy balance models for purposes such as determining stream temperatures to assess aquatic habitat suitability. In long enough model reaches, less accurate downstream temperatures due to the exclusion of the diurnal discharge signal may eventually have a large enough effect on the overall average temperature of the model reach to alter predictions of average groundwater inflow over the reach. Lastly, incorporating fluctuations in diurnally fluctuating stream discharge may improve estimates of the groundwater temperature entering the reach. Because there are many locations where the daily variations in stream discharge are poorly constrained, the results from these model simulations

provide a means of assessing when these diurnal variations are important for simulating temperatures in a particular stream and therefore understanding the relative importance of different processes on stream temperatures.

References

Baker, E., Lautz, L., Kelleher, C., & McKenzie, J. (2018). The importance of incorporating diurnally fluctuating stream discharge in stream temperature energy balance models, *HydroShare*, https://doi.org/10.4211/hsb96d3056d0ff488c87e6659c871cf05c

Boyd, M., & Kasper, B. (2003). Analytical methods for dynamic open channel heat and mass transfer: Methodology for the heat source model Version 7.0, 204. Retrieved from http://www.deq.state.or.us/wq/TMDLs/tools.htm

Cassie, D. (2006). The thermal regime of rivers: A review. *Freshwater Biology*, 51, 1389–1406. https://doi.org/10.1111/j.1365-2427.2006.01567.x

Deutscher, J., Kupec, P., Dundek, P., Holík, L., Machala, M., & Urban, J. (2016). Diurnal dynamics of streamflow in an upland forested microwatershed during short precipitation-free periods is altered by tree sap flow. *Hydrological Processes*, 30, 2042–2049. https://doi.org/10.1002/hyp.10771

Dugdale, S. J. (2016). A practitioner's guide to thermal infrared remote sensing of rivers and streams: Recent advances, precautions and considerations. Wiley Interdisciplinary Reviews: *WIREs Water*, 3, 251–268. https://doi.org/10.1002/wat2.1135

Dugdale, S. J., Hannah, D. M., & Malcolm, I. A. (2017). River temperature modelling: A review of process-based approaches and future directions. *Earth-Science Reviews*, 175, 97–113. https://doi.org/10.1016/j.earscirev.2017.10.009

Glose, A. M., Lautz, L. K., & Baker, E. A. (2017). Stream heat budget modeling with HFLUX:

model development, verification, and applications across contrasting sites and seasons. *Environmental Modeling & Software*, 92, 213–228. https://doi.org/10.1016/j.envsoft.2017.02.021

Gordon, R. P., Lautz, L. K., McKenzie, J. M., Mark, B. G., Chavez, D., & Baraer, M. (2015). Sources and pathways of stream generation in tropical proglacial valleys of the Cordillera Blanca, Peru. *Journal of Hydrology*, 522, 628–644. https://doi.org/10.1016/j.jhydrol.2015.01.013

Greimel, F., Zeiringer, B., Holler, N., Grun, B., Godina, R., & Schmutz, S. (2016). A method to detect and characterize sub-daily flow fluctuations. *Hydrological Processes*, 30, 2063–2070. https://doi.org/10.1002/hyp.10773

Gu, R., Montgomery, S., & Austin, T. A. (1998). Quantifying the effects of stream discharge on summer river temperature. *Hydrological Sciences Journal*, 43(6), 885–904. https://doi.org/10.1080/02626669809492185

Heavilin, J. E., & Neilson, B. T. (2012a). An analytical solution to main channel heat transport with surface heat flux. *Advances in Water Resources*, 47, 67–75. https://doi.org/10.1016/j.advwatres.2012.06.006

Heavilin, J. E., & Neilson, B. T. (2012b). Approximation of inverse Laplace transform solution to heat transport in a stream system. *Water Resources Research*, 48(9), 1–6. https://doi.org/10.1029/2012WR012329

Kaser, G., Ames, A., & Zamora, M. (1990). Glacier fluctuations and climate in the Cordillera Blanca, Peru. *Annals of Glaciology*, 14, 136–140.

Kim, K. S., & Chapra, S. C. (1997). Temperature model for highly transient shallow streams. *Journal of Hydraulic Engineering*, 123, 30–40.

King, T. V., Neilson, B. T., Overbeck, L. D., & Kane, D. L. (2016). Water temperature controls in low arctic rivers. *Water Resources Research*, 52, 1–19. https://doi.org/10.1002/2015WR017965

Kurylyk, B. L., Moore, R. D., & MacQuarrie, K. T. B. (2015). Scientific briefing: Quantifying streambed heat advection associated with groundwater-surface water interactions. *Hydrological Processes*, 30, 987–992. https://doi.org/10.1002/hyp.10709

Loheide, S. P., & Gorelick, S. M. (2006). Quantifying stream–aquifer interactions through the analysis of remotely sensed thermographic Profiles and in situ temperature histories. *Environmental Science and Technology*, 40(10), 3336–3341. https://doi.org/10.1021/es0522074

Loheide, S. P. II, & Lundequist, J. D. (2009). Snowmelt-induced diel fluxes through the hyporheic zone. *Water Resources Research*, 45, W07404. https://doi.org/10.1029/2008WR007329

Mark, B. G., & Seltzer, G. O. (2003). Tropical glacier meltwater contribution to stream discharge: A case study in the Cordillera Blanca, Peru. *Journal of Glaciology*, 49(165), 271–281. https://doi.org/10.3189/172756503781830746

Pianosi, F., Sarrazin, F., & Wagener, T. (2015). A Matlab toolbox for global sensitivity analysis. *Environmental Modelling and Software*, 70, 80–85. https://doi.org/10.1016/j.envsoft.2015.04.009

Schmadel, N. M., Neilson, B. T., & Heavilin, J. E. (2015). Spatial considerations of stream hydraulics in reach scale temperature modeling. *Water Resources Research*, 51, 5566–5581. https://doi.org/10.1002/2015WR016931

Smokorowski, K. E. (2010). Effects of experimental ramping rate on the invertebrate community

of a regulated river. Proceedings of the Colorado River Basin Science and Resource Management Symposium, November 18–20, 2008, Scottsdale, Arizona: *U.S. Geological Survey Scientific Investigations Report* 2010–5135, 372 p.

Somers, L. D., Gordon, R. P., McKenzie, J. M., Lautz, L. K., Wigmore, O., Glose, A. M., … Condom, T. (2016). Quantifying groundwater-surface water interactions in a proglacial v alley, Cordillera Blanca, Peru. *Hydrological Processes*, 30(17), 2915–2929. https://doi.org/10.1002/hyp.10912

Suarez, W., Chevallier, P., Pouyaud, B., & Lopez, P. (2008). Modelling the water balance in the glacierized Parón Lake basin (White Cordillera, Peru). *Hydrological Sciences Journal*, 53(1), 266–277. https://doi.org/10.1623/hysj.53.1.266

Wagener, T., & Kollat, J. (2007). Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox. *Environmental Modelling and Software*, 22(7), 1021–1033. https://doi.org/10.1016/j.envsoft.2006.06.017

Webb, B. W., & Zhang, Y. (1997). Spatial and seasonal variability in the components of the river heat budget. *Hydrological Processes*, 11, 79–101.

Webb, B. W., & Zhang, Y. (1999). Water temperatures and heat budgets in Dorset chalk water courses. *Hydrological Processes*, 13, 309–321.

Westhoff, M. C., Savenije, H. H. G., Luxemburg, W. M. J., Stelling, G. S., van de Giesen, N. C., Selker, J. S., … Uhlenbrook, S. (2007). A distributed stream temperature model using high resolution temperature observations. *Hydrology and Earth System Sciences*, 11, 1469–1480.

Wigmore, O., & Mark, B. (2017). High altitude kite mapping: Evaluation of kite aerial

photography (KAP) and structure from motion digital elevation models in the Peruvian Andes. *International Journal of Remote Sensing*, 1–21. https://doi.org/10.1080/01431161.2017.1387312

Wondzell, S. M., Gooseff, M. N., & McGlynn, B. L. (2007). Flow velocity and the hydrologic behavior of streams during baseflow. *Geophysical Research Letters*, 34, L24404, https://doi.org/10.1029/2007GL031256

Wondzell, S. M., Gooseff, M. N., & McGlynn, B. L. (2010). An analysis of alternative conceptual models relating hyporheic exchange flow to diel fluctuations in discharge during baseflow recession. *Hydrological Processes*, 24, 686–694. https://doi.org/10.1002/hyp.7507

Yearsley, J. R. (2009). A semi-Lagrangian water temperature model for advection-dominated river systems. *Water Resources Research*, 45(12), 1–19. https://doi.org/10.1029/2008WR007629

Younus, M., Hondzo, M., & Engel, B. (2000). Stream temperature dynamics in upland agricultural watersheds. *Journal of Environmental Engineering*, 126(6), 618–628. https://doi.org/10.1061/(ASCE)0733-9372(2000)126:6(518

Tables

Table 1. Ranges of model input parameters explored during the sets of simulations.

Input	Range	Justification
Timing of peak discharge	0:00 to 23:59	Whole 24 hr cycle (Wondzell et al., 2007)
Amplitude of discharge fluctuation	0% to 50%	Upper limit of diurnal amplitude values for natural streams (e.g. Deutscher et al., 2016)
Mean discharge	0.6 to 1.0 m^3s^{-1} (-20% to +33.3%)	Brackets observed range of Quilcay discharge during study period ±0.1 m^3s^{-1}
Bank angle (2θ)	10° (deep and narrow channel) to 189° (shallow and wide channel)	Complete range of plausible bank angles
Groundwater Inflow Rate	-20% to +20% (-150 to +150 ls^{-1}) or -20% to +80% (-150 to + 600 ls^{-1})	Range in Somers et al. (2016)[a]
Groundwater Temperature	4 to 18°C for Figure 8 output 10.6°C for Figure 5c, 7 & 9 output	Range in Somers et al. (2016) Average from the 5 wells

[a]The groundwater inflow rate varied from −20% to +20% of the stream discharge rate for the Monte Carlo and nonrandom simulations where amplitude, timing, and groundwater flux were varied. The groundwater inflow rate was varied from −20% to +80% for the Monte Carlo simulation where the groundwater temperature and inflow rate were varied to be consistent with Somers et al. (2016). A negative value indicates a losing stream; no groundwater inflow with stream water exiting the stream.

Table 2. Results of the base case model simulation using constant versus observed streamflow hydrographs for the downstream reach (5 sensors over 94.3 m).

Error Metric	Constant Q	Observed Q (diurnal pattern)
RMSE (°C)	0.27	0.26[a]
Mean Difference – Peak Temperatures (°C)	-0.07	-0.06[a]
Mean Difference – Minimum Temperatures (°C)	0.34	0.33[a]
Mean Difference – Mid - Morning Temperatures (°C)	-0.15[a]	-0.17
Mean Difference – Early Evening Temperatures (°C)	0.25	0.24[a]

Note. The mean temperature difference is expressed as the average of the measured stream temperature minus the modelled temperature, at each time over all of the days. RMSE, root mean square error.

[a]Indicates the smaller error.

Figures

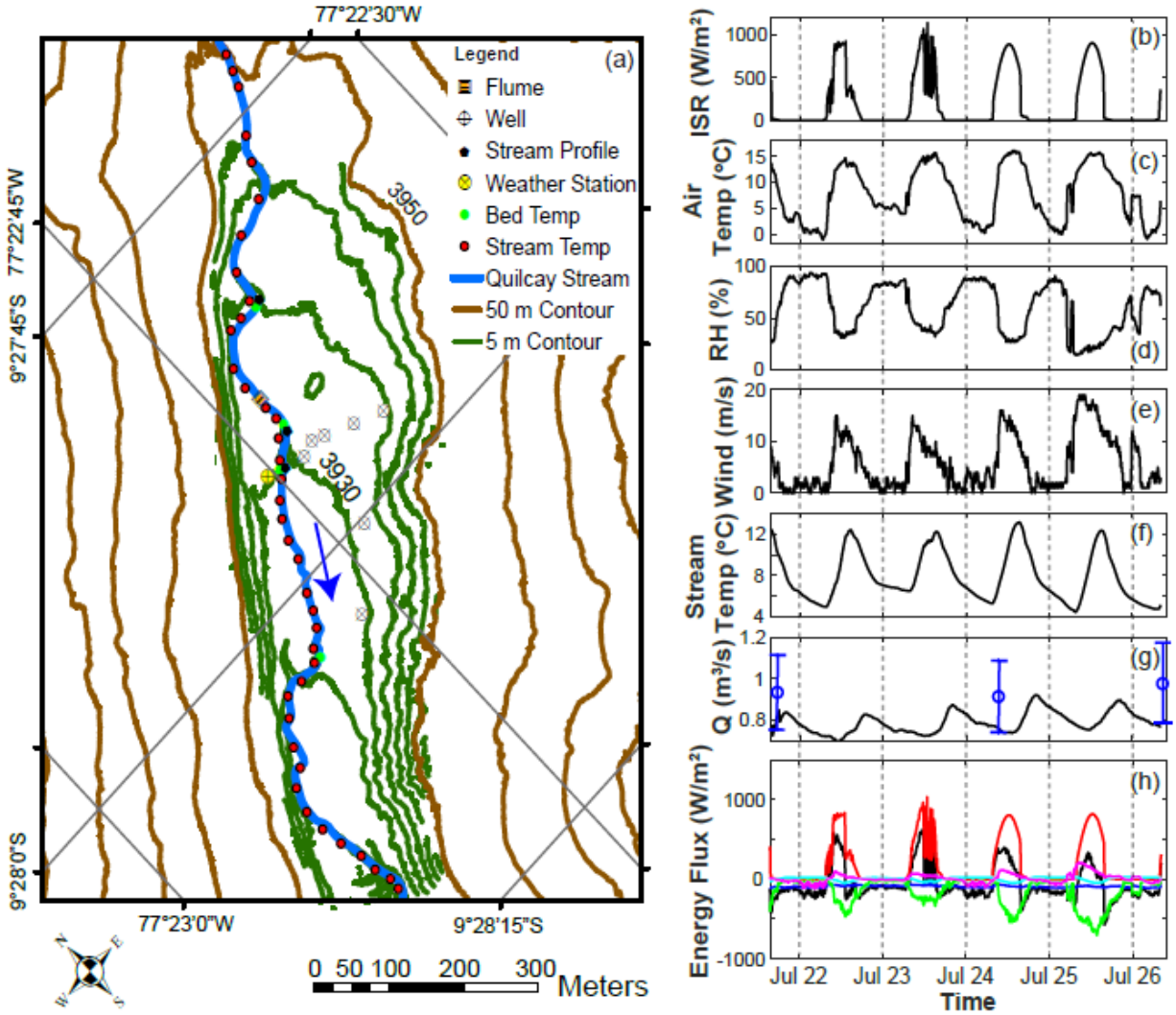

Figure 1. (a) Locations of data collected within Quilcayhuanca Valley. The blue arrow indicates flow direction. Measured data include the (b) incoming shortwave radiation, (c) air temperature, (d) relative humidity, (e) wind speed, (f) mean stream temperature at all sites, and (g) stream discharge, where the error bars represent the uncertainty (±20%) in the manual discharge measurements. (h) The calculated heat fluxes for the base case model using the observed stream discharge. The heat fluxes depicted include the total heat flux (black), net shortwave radiation (red), longwave radiation (blue), latent heat flux (green), streambed conduction (teal), and sensible heat flux (magenta).

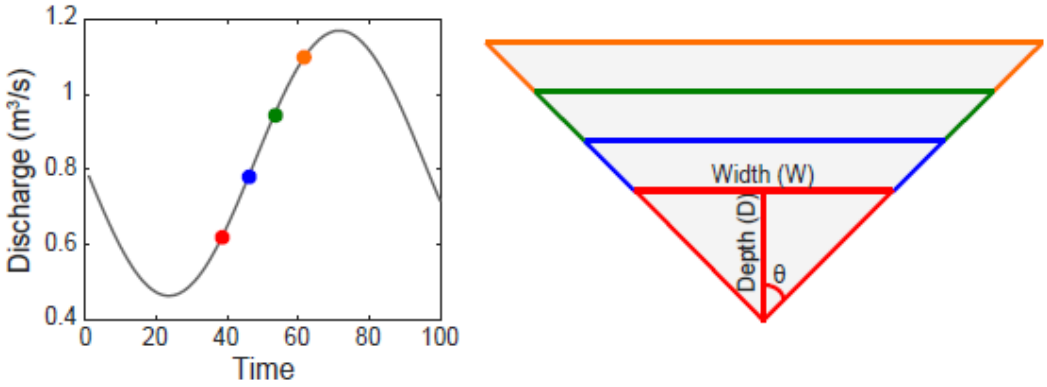

Figure 2. As the discharge changes, the depth (D) and width (W) of the stream channel change by holding the angle (θ), slope (s), and Manning's roughness coefficient (n) constant for any given location. The bank angle is equal to 2θ.

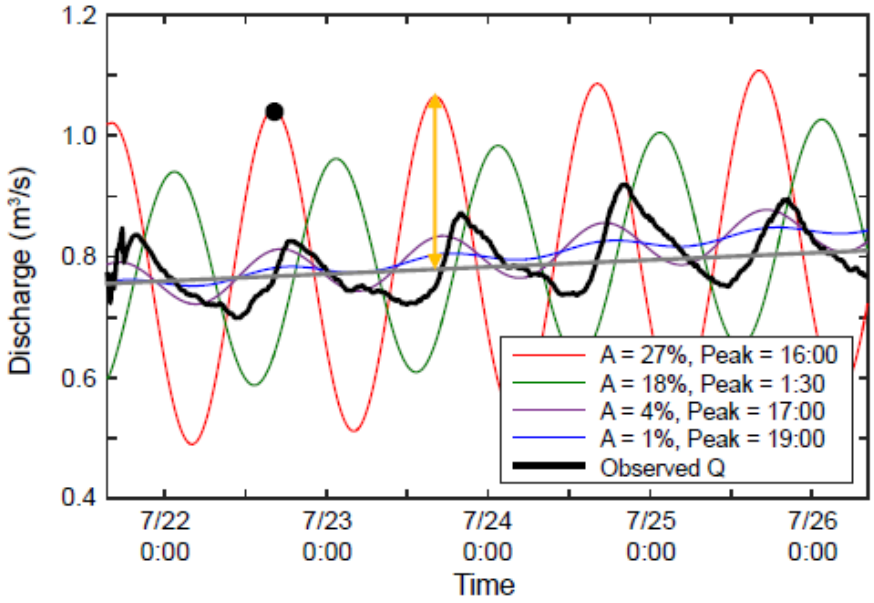

Figure 3. Diurnal discharge fluctuations can occur from a variety of causes. In the Quilcay stream (black line), diurnal fluctuations occur due to glacial melt. The other lines show examples of simulated diurnal discharge fluctuations. The average discharge through time is the grey line about which the diurnal signal fluctuates, the yellow arrow indicates the amplitude of the signal, and the phase or timing describes when the diurnal discharge fluctuation either reaches a local peak (black dot) or minimum.

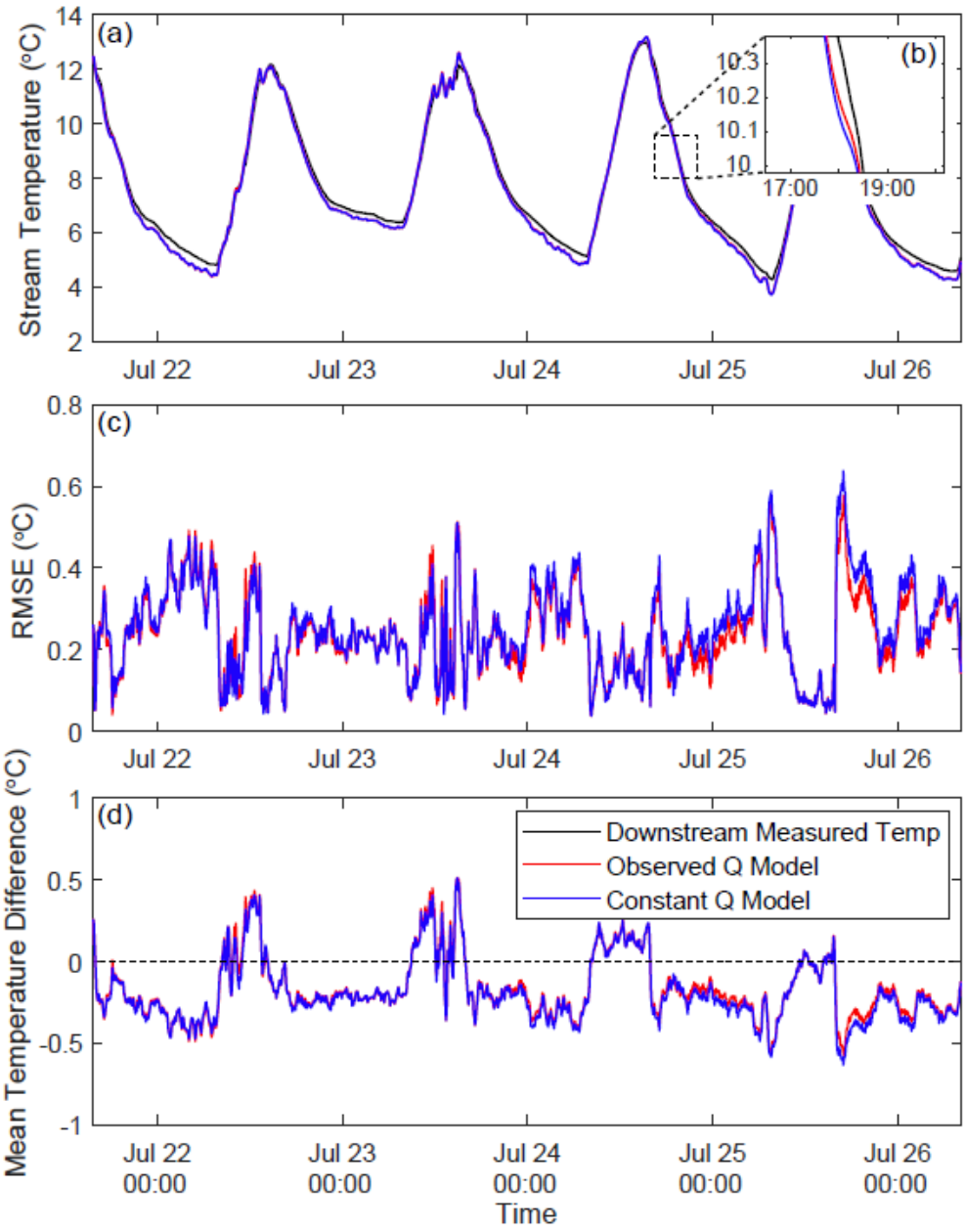

Figure 4. (a) Average measured downstream temperatures through time and HFLUX model temperature results using observed and constant discharge signals, assuming no groundwater influx. (b) Comparing decreasing stream temperature results. (c) The root mean square error (RMSE) of the downstream temperature through time. (d) The mean downstream temperature difference through time where positive values indicate modelled temperatures are warmer than measured stream temperatures.

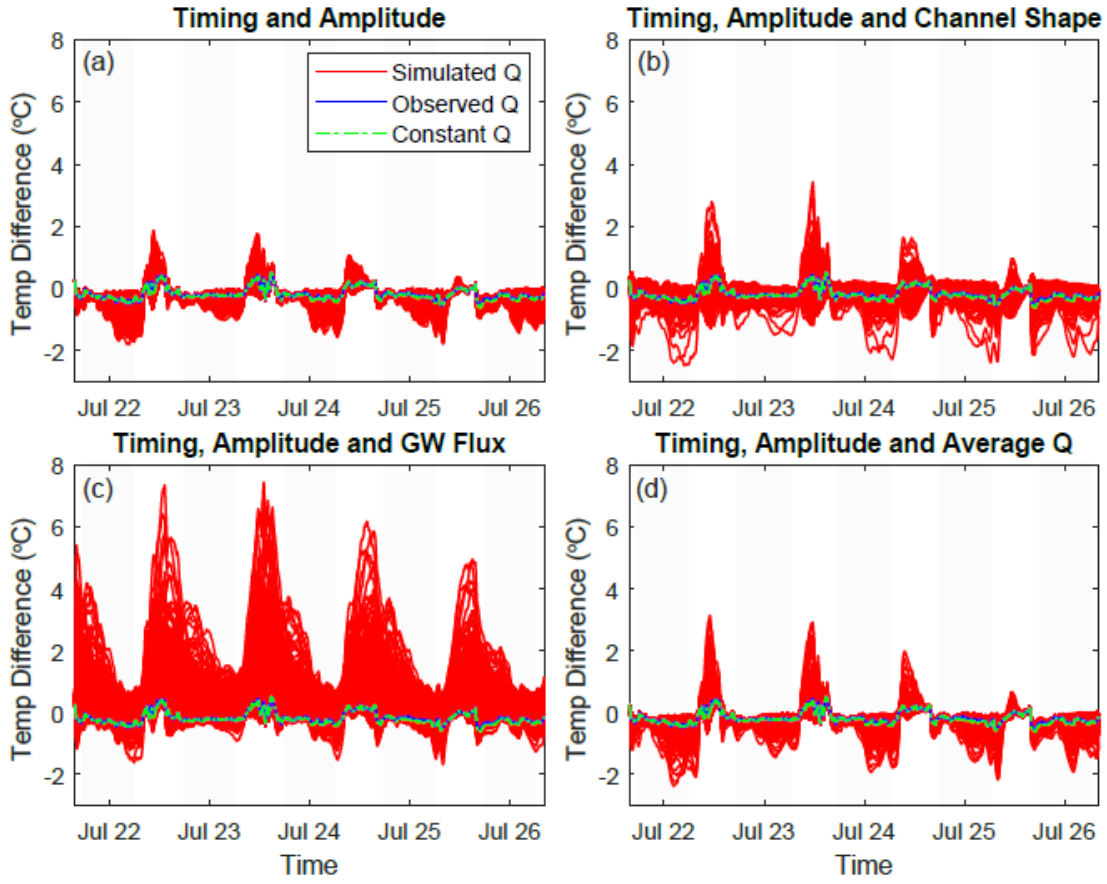

Figure 5. Results from Monte Carlo simulations (red) showing the temperature difference between the modelled and measured stream temperatures through time, with grey areas indicating night-time periods. The blue line is the model run when the observed discharge is used, while the green line is when constant (750 Ls^{-1}) discharge is used in the model. The blue and green lines essentially overlap. Positive values indicate modelled temperatures are warmer than measured temperatures. (a) Set of models where the timing and amplitude of the stream discharge were varied. (b) Set of models where the timing and amplitude of the stream discharge were varied along with the channel shape. (c) Set of models where the timing and amplitude of the stream discharge were varied along with the groundwater flux (10.6°C groundwater). (d) Set of models where the timing, amplitude, and mean through time of the stream discharge were varied.

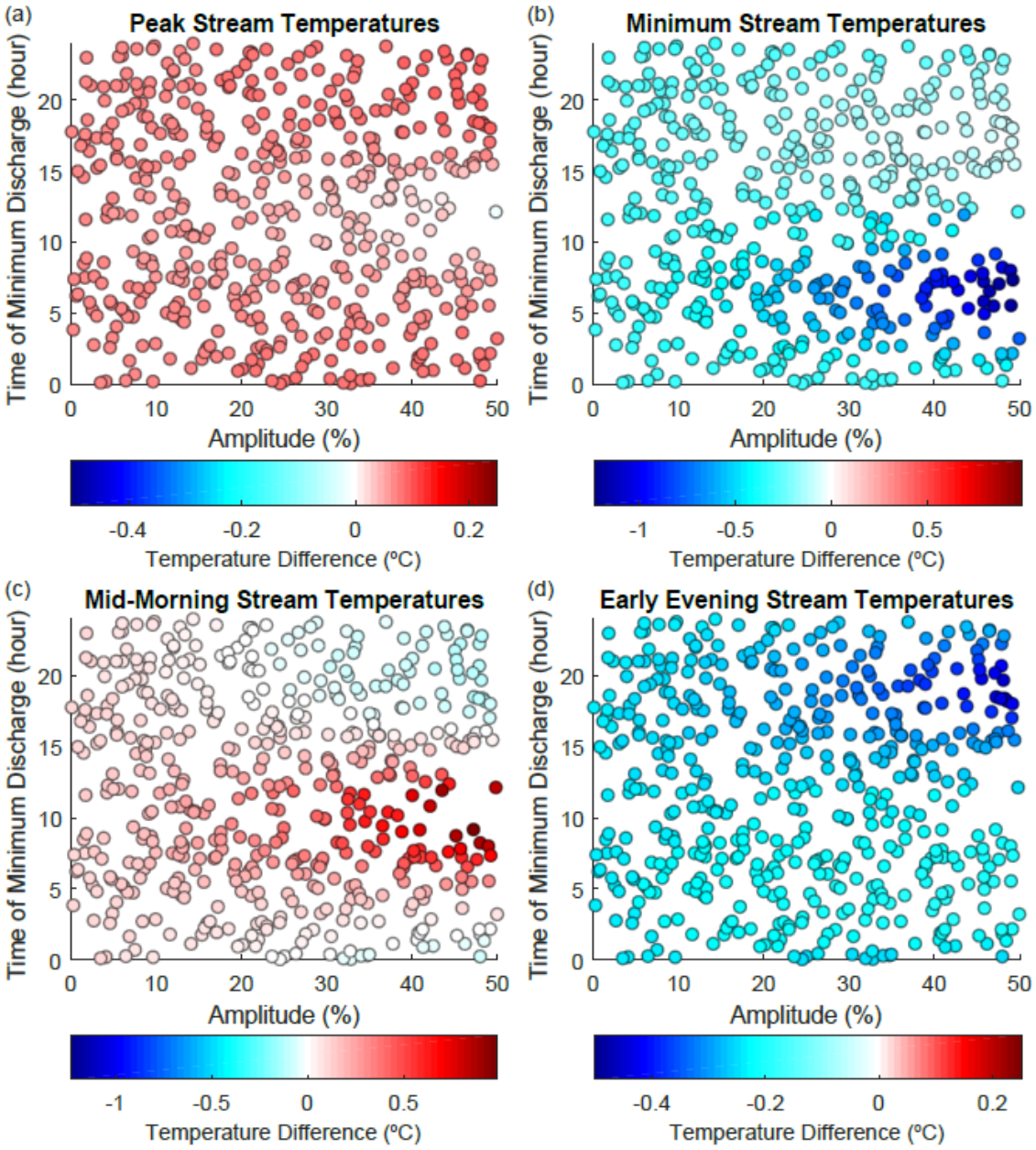

Figure 6. Output from 500 Monte Carlo model runs varying the timing of the minimum discharge and the amplitude of the diurnal discharge fluctuations. The average temperature difference between the modelled and observed stream temperatures was calculated at the (a) peak stream temperature, (b) minimum stream temperature, (c) midpoint of the increasing stream temperature, and (d) midpoint of the decreasing stream temperature of each day. Temperature differences range from −1.2°C to 1.0°C.

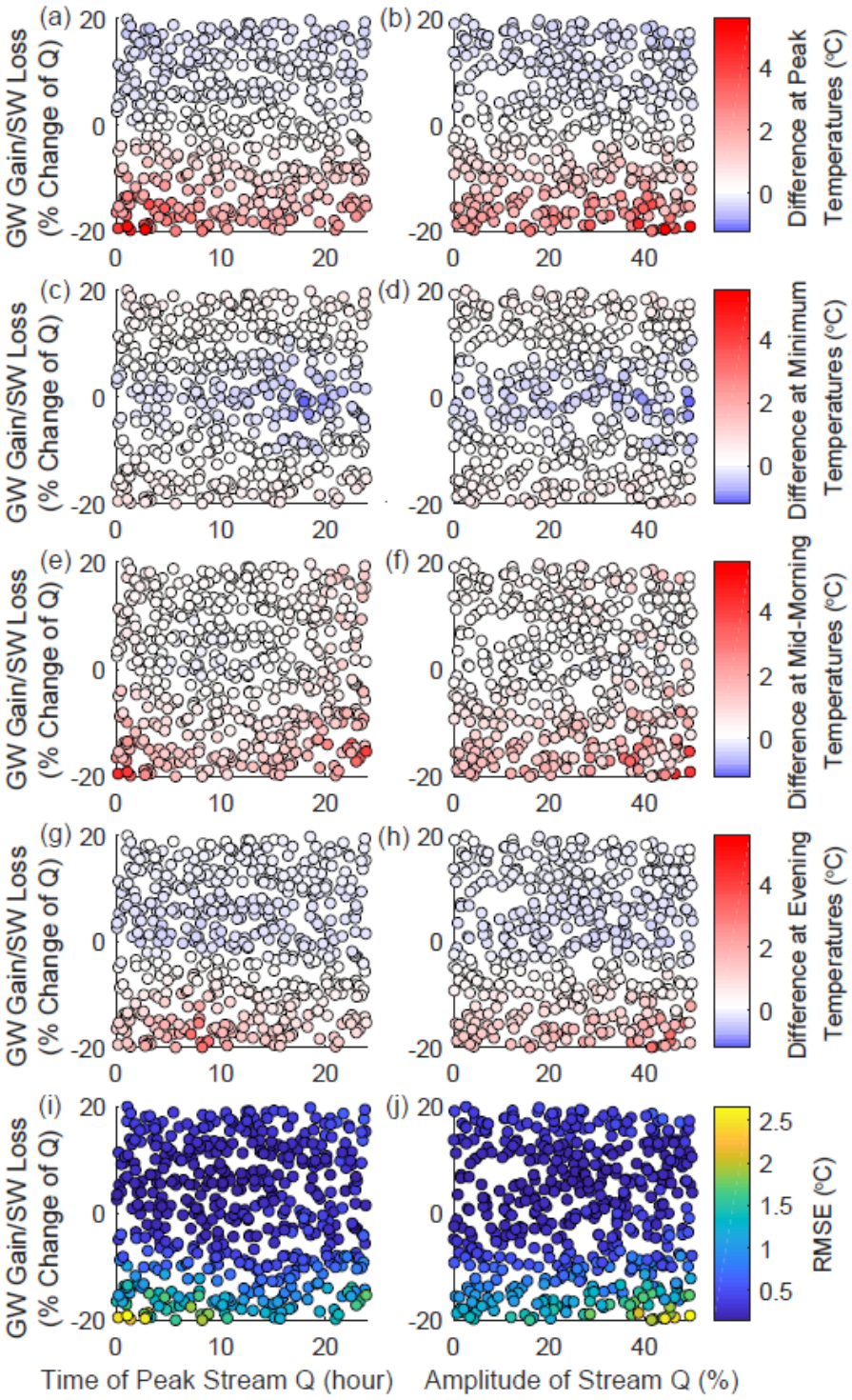

Figure 7. Output from 500 Monte Carlo simulations varying the diurnal stream discharge timing and amplitude, and the change in streamflow due to groundwater-surface water exchange. A negative change in discharge (SW loss) represents a losing stream, whereas a positive change in

discharge (GW gain) indicates a gaining stream due to groundwater influx. Model behavior was examined using downstream temperature differences at the (a,b) peak stream temperatures, (c,d) minimum stream temperatures, (e,f) midpoint of the increasing morning temperatures, (g,h) midpoint of the decreasing evening temperatures, and (i,j) using root mean square error (RMSE).

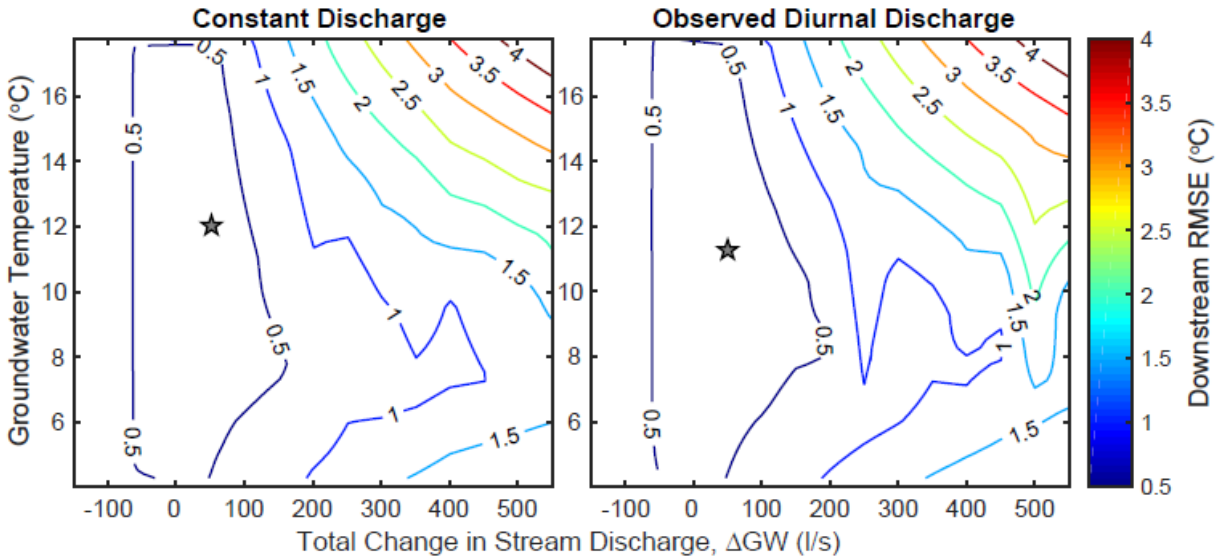

Figure 8. Error metrics used to estimate the optimal groundwater temperature and groundwater flux over the ~1.2-km model reach for both the (a) constant discharge and (b) diurnally varying discharge scenarios. The grey star indicates the optimal groundwater flux and groundwater temperature for the reach. The average observed groundwater temperature was 10.6°C in the five study wells. Contour lines are based on 500 Monte Carlo simulations. Contour lines occur at intervals of 0.5°C.

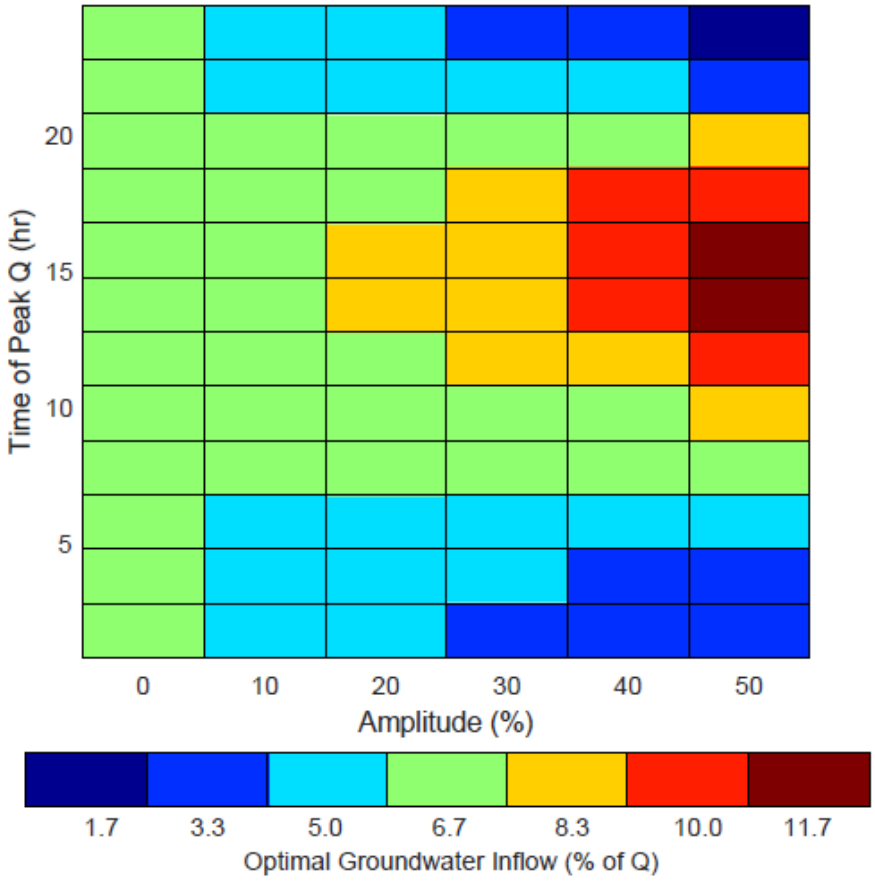

Figure 9. Optimized groundwater contribution to the stream for each timing and amplitude combination of diurnal fluctuation in stream discharge. The optimized value was obtained from the model run with the lowest downstream RMSE using a groundwater temperature of 10.6°C. The optimized groundwater contribution is in percent of the initial upstream discharge of 750 L s^{-1}. A contribution of 1.7% represents a groundwater inflow of 12.5 L s^{-1} over the length of the ~1.2-km study reach, whereas a contribution of 11.7% represents an influx of 87.5 L s^{-1} over the reach.

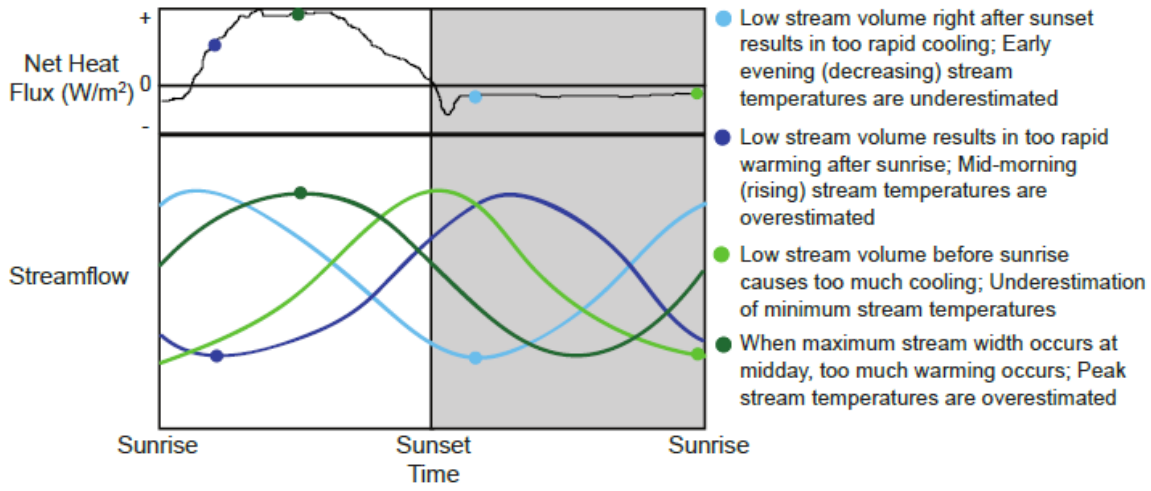

Figure 10. Summary of interactions between diurnal discharge timing, heat fluxes, and resulting model error in stream temperature. Impacts are greatest when the amplitude of fluctuations in discharge is at a maximum.

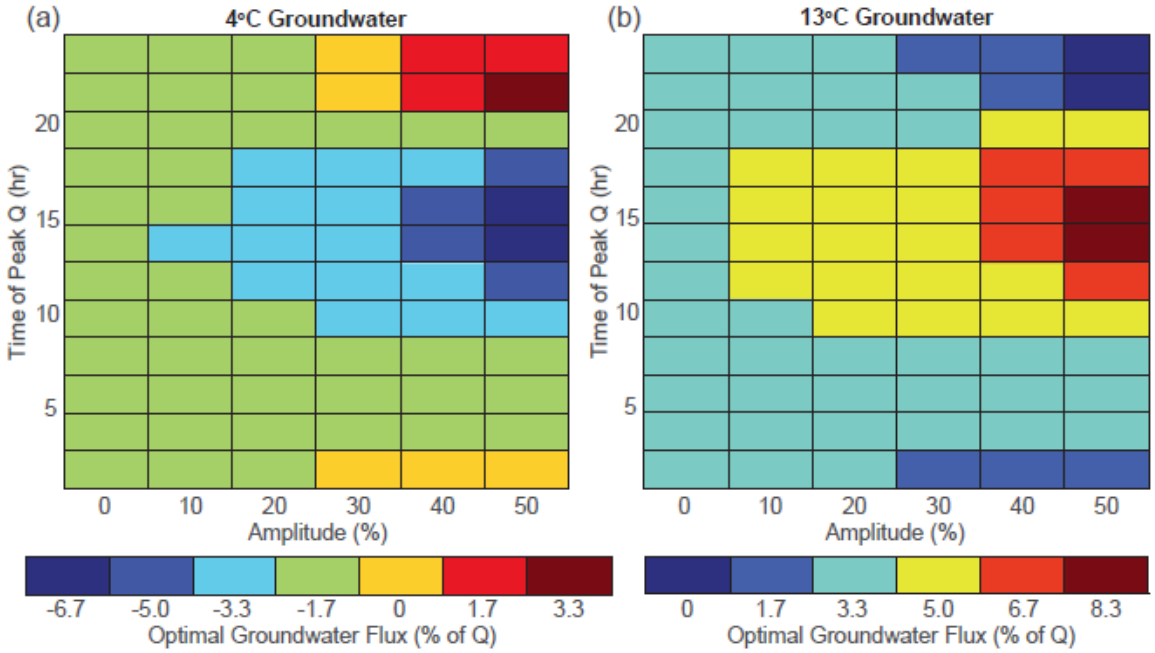

Figure S1. Optimized groundwater flux to stream for each timing and amplitude combination of diurnal fluctuation in stream discharge using groundwater temperatures of (a) 4°C and (b) 13°C. The optimized flux value was obtained from the model run with the lowest downstream RMSE. A negative flux value indicates a losing stream; no groundwater inflow with stream water exiting the stream.

Chapter 2

Improving the Accuracy of Time-Lapse Thermal Infrared Imaging for Hydrologic Applications

Published as:

Baker, E.A., L.K. Lautz, J.M. McKenzie, C. Aubry-Wake. 2019. Improving the accuracy of time-lapse thermal infrared imaging for hydrologic applications, *Journal of Hydrology*, 571, 60-70. https://doi.org/10.1016/j.jhydrol.2019.01.053

Abstract

In recent years, thermal infrared (TIR) cameras have improved in resolution and accuracy while their cost has declined. By deploying a ground-based TIR camera to collect time-lapse images, it is now possible to acquire high-resolution stream temperatures through both space and time. However, while ground-based TIR is useful for qualitatively identifying stream temperature differences, acquisition of absolute stream temperatures remains difficult due to interference from reflected radiation. Therefore, improved correction approaches are still needed to extract absolute stream temperatures from ground-based, time-lapse TIR imagery. Using >1100 TIR images acquired every 10 minutes during two field seasons, we assess two methods for correcting time-lapse, ground-based TIR stream temperature data: (1) an analytical method derived from the literature that corrects for atmospheric transmissivity, reflected temperatures and water surface emissivity, which did not improve TIR temperature accuracy, and (2) an empirical approach that uses an offset correction created from in-stream control point temperatures, which reduced the mean absolute temperature difference between the TIR and in situ stream temperatures. Examination of the analytical method revealed its sensitivity to reflected temperatures from the surrounding environment, a problem that is particularly pronounced in ground-based TIR imagery due to the lower stream emissivity at more oblique viewing angles. Since reflected temperatures and stream surface emissivity can be difficult to quantify and are misrepresented in previous hydrologic literature, the empirical correction approach offers an alternative method for extracting absolute stream temperatures from ground-based TIR imagery affected by these factors.

1 Introduction

Satellites, airborne sensors, and ground-based sensors are increasingly used to collect thermal infrared (TIR) data for environmental applications. In hydrology, thermal data can be used to reveal stream temperature heterogeneities caused by groundwater inflows and seeps, hyporheic exchange, and geothermal inputs (e.g. Torgersen et al., 2001; Handcock et al., 2012; Hare et al., 2015; Eschbach et al., 2016; Mundy et al., 2017). However, the spatial resolution of satellite TIR data is too coarse for all but the widest rivers (Handcock et al., 2006; Schuetz and Weiler, 2011), airborne TIR data collection is inhibited by cost (Vatland et al., 2015), and UAV platforms are limited by battery life. Additionally, satellite data are infrequent, and taking high frequency airborne surveys from UAVs or helicopters is difficult, preventing the comparison of TIR imagery to output from stream temperature models through time. Traditional in-stream temperature sensors such as fibre optic technology and data logging thermocouples have limited spatial resolution, particularly in 2D plan view (Torgersen et al., 2001; Cardenas et al., 2011). Time-lapse, ground-based TIR photography can potentially acquire large quantities of spatially and temporally detailed data, revealing heterogeneities in river temperatures as never before. While this data can be useful for qualitatively identifying stream temperature heterogeneities, acquisition of accurate absolute stream temperatures remains difficult for ground-based TIR imagery due to interference from reflected radiation.

While many studies have used satellite, airborne, and ground-based TIR sensors for instantaneous observations of relative stream temperature differences, few have attempted to use time-lapse, ground based TIR to determine absolute temperatures through time. Three recent studies used this method to investigate environmental processes, but none address the issue of reflections, which are more pronounced when using ground-based sensors due to the viewing

angle. Aubry-Wake et al. (2015, 2018) used time-lapse, ground-based infrared imaging to measure glacial surface temperatures and calculate the glacial energy budget. However, an ~6°C offset correction was applied to the TIR images after analytically correcting glacier temperatures. Cardenas et al. (2014) compared time-lapse, ground-based TIR imagery to modeled stream temperatures and determined warm inflows were needed to model the observed TIR stream temperature patterns. However, only one in-stream control point that could not be directly viewed by the camera was reported (Cardenas et al., 2014). Our paper builds on these prior time-lapse, ground-based TIR studies by further investigating some strengths and limitations of this type of data, particularly its sensitivity to reflected temperatures from the surrounding environment. We assess two TIR correction methods (analytical and empirical) and evaluate the effectiveness of time-lapse, ground-based TIR imagery for observing absolute stream temperatures. Drawing on best practices for TIR sensing of sea surface temperatures (SST), we propose methodological changes to improve the accuracy of time-lapse, ground-based TIR stream temperatures when using the analytical method.

2 Methodology

2.1 Thermal Infrared Science

TIR cameras record the intensity of surface radiation emitted by the top 0.1 mm of objects between frequencies of 7.5 and 14 μm and relate it to temperature using Planck's Law. TIR data must be corrected to isolate the surface temperature ($T_{surface}$) of an object because the total measured radiation (L_{Tmeas}) comes from the target object ($L_{Tsurface}$), atmospheric radiation (L_{Tatm}), and reflected radiation (L_{Trefl}), according to the equation:

$$L_{(T_{meas})} = L_{T_{surface}}\varepsilon\tau + L_{T_{atm}}(1-\tau) + L_{T_{refl}}(1-\varepsilon)\tau \qquad \text{[Eq. 1]}$$

where L is the radiant energy measured in $Wsr^{-1}\ m^{-3}$ (Banks et al., 1996; Cardenas et al., 2011; Handcock et al., 2012; Aubry-Wake et al., 2015). This equation corrects for the target object's surface emissivity (ε), atmospheric transmissivity (τ), and reflected and atmospheric radiation (L_{Trefl}, L_{Tatm}) (e.g. Torgersen et al., 2001; Cardenas et al., 2014; Aubry-Wake et al., 2015).

Atmospheric transmissivity (τ) quantifies the attenuation of a target object's radiation due to scattering and absorption by the atmospheric column (e.g. Handcock et al., 2012). The transmissivity is affected by the distance between the camera and target object, relative humidity, and atmospheric temperature (e.g. Atwell et al., 1971; Torgersen et al., 2001; Kay et al., 2005). Previous studies correct for atmospheric radiation and transmissivity either using FLIR software which applies proprietary empirical formulas based on the LOWTRAN atmospheric model (e.g. Bingham et al., 2012; Cardenas et al., 2014; Aubry-Wake et al., 2015) or using MODTRAN, LOWTRAN or other radiative transfer models (e.g. Torgersen et al., 2001; Kay et al., 2005; Fricke and Baschek, 2015).

Emissivity (ε) is the ratio of radiation an object emits compared to a black body at the same temperature (Buettner and Kern, 1965); for water, emissivity typically ranges from 0.95 to 0.99 when viewing the surface at nadir and can be affected by viewing angle, turbulence/surface roughness (e.g. due to high winds), salinity, turbidity, and other factors (Zappa and Jessup, 1998; Torgersen et al., 2001; Jessup and Branch, 2008; Cardenas et al., 2011). As surface emissivity decreases, reflectivity increases, according to the equation:

$$\varepsilon + r + \tau_{obj} = 1 \qquad \text{[Eq. 2]}$$

where ε is the object's emissivity, r is the object's reflectivity, and τ_{obj} is the object's transmissivity. This equation reduces to $\varepsilon + r = 1$ for water since it cannot transmit infrared radiation (Buettner and Kern, 1965; Sidran, 1981; Salisbury and D'Aria, 1992). As surface

emissivity decreases, a greater proportion of the radiation from the object is radiation reflected from the surrounding environment, interfering with the desired thermal signal (Gillespie et al., 1998; Anderson et al., 1995; Torgersen et al., 2001; Kay et al., 2005; Puleo et al., 2012). Therefore, TIR images should be taken from near nadir to maintain high emissivity and reduce reflected radiation (Torgersen et al., 2001; Dugdale, 2016).

2.2 Study Site

TIR images were acquired of the Quilcay Stream in the Quilcayhuanca Valley located in the Cordillera Blanca, Peru (9.4656°S, 77.3792°W). The site experiences diurnal fluctuations in kinetic stream temperature, relative humidity, and wind speed; these fluctuations are similar both from day to day within the dry season, and similar from one dry season to another (Fig. 1). Peak incoming solar radiation often exceeds 900 W/m^2. The land surface of the valley bottom is ~160 to 220 m across at the camera location, at an elevation of ~3930m asl. Steep, granodiorite valley walls extend to an elevation of over 5200m asl. The Quilcay Stream, containing a gravel bar, flows down the valley, with the study reach traversing a meadow containing short grazed grasses and no overhanging vegetation. Stream width in the study reach varies from ~3 to 14 m, with a maximum depth of ~0.47 m. Numerous small springs and tributaries flow into the stream.

2.3 Field Methods

Time-lapse, ground-based TIR images of the Quilcay Stream were acquired using a Jenoptik VarioCam high-definition TIR camera (Table S1) during two field periods (July 2015, August 2016; Table 1). TIR and visual images of the stream were acquired every 10 min (Fig. 2a–d). During 2015, we positioned the camera on the south-east valley wall ~110m above the stream. We recorded 707 time-lapse images of a ~500m stream reach, with a spatial resolution of ~50 cm (Fig. 2c). The image center was viewed at ~77° from vertical, with the upstream

edge viewed at ~82° and the downstream edge at ~72°. During 2016 we collected 441 images of a sub-reach of the 2015 reach with a resolution of ~10 cm. We deployed the camera on the north-west valley wall ~51m above the stream with a view containing an ~80m stream reach (Fig. 2d). The viewing angle of the camera was ~70° from vertical. Failure of the TIR camera's backup battery resulted in a data gap during the 2016 field period (Table 1).

Thermochron iButton sensors (Table S1) directly recorded in-stream kinetic water temperatures every 5 min, serving as control points for the TIR stream temperatures. In 2015, these in-stream sensors were deployed at 17 locations (Fig. 3a) and in 2016 they were deployed at 13 locations (Fig. 3b). In 2015 we also installed three in-stream stakes, each with iButton sensors at three depths, to obtain water column temperature profiles to confirm the stream was thermally well mixed, ensuring TIR temperatures are representative of water column temperatures. iButton sensors were also installed in the surrounding environment during both field periods to record ground temperatures every 5 min. Ground control points were buried about a centimeter below the surface in a variety of surface types including within wet areas of the meadow, dry areas of the meadow, and the gravel bar. Ten ground control points were deployed in 2015 and three were deployed in 2016. In both years, meteorological data (air temperature, relative humidity, wind speed, and incoming solar radiation) were recorded every 10 min using a Vantage Pro2 weather station centrally located in the study reach (Fig. 3; Fig. 1). At the camera location, a Lascar sensor also recorded air temperature and relative humidity every 5 min.

2.4 Image Processing and Corrections

Two methods were assessed for correcting the radiant stream temperatures (Fig. 4). The first was the analytical method which used Planck's Law and Eq. (1) to correct radiant

temperatures for atmospheric transmissivity (τ), reflected temperature (T_{refl}), and stream surface emissivity (ε) (Cardenas et al., 2011; Cardenas et al., 2014; Aubry-Wake et al., 2015). The second was an empirical method which used an offset correction to account for errors due to reflected radiation. Prior to the application of either correction approach, images were aligned using functions within Matlab's Image Processing Toolbox (Although the camera was securely anchored, the field of view occasionally shifted slightly while changing batteries and downloading data). Blurred images (due to camera vibrations from the wind) were removed (3 removed in 2015, 5 removed in 2016). The downstream distance of each pixel was measured in ArcMap. Stream reaches with widths spanned by fewer than 8 pixels were excluded from the analysis to minimize thermal contamination from the stream banks (Torgersen et al., 2001; Handcock et al., 2006); 7 and 10 control points from 2015 and 2016 were located in reaches with enough pixel coverage for analysis.

For the analytical correction method, Planck's Law was used to calculate the radiance measured by the camera (Atwell et al., 1971; Torgersen et al., 2001). The measured radiance was corrected for stream surface emissivity, atmospheric transmissivity, and reflected and atmospheric radiation using Eq. (1) (Cardenas et al., 2011; Cardenas et al., 2014; Aubry-Wake et al., 2015). Reflected and atmospheric radiation were calculated using Planck's Law and air temperatures recorded at the camera and weather station (Aubry-Wake et al., 2015). Finally, corrected radiances were converted back to temperatures (Aubry-Wake et al., 2015). τ values were determined using FLIR Tools for a range of air temperature, relative humidity, and distance combinations and used within Eq. (1). The air temperature and relative humidity for the atmospheric transmissivity correction were obtained from the weather station. The distance to the ground surface from the TIR camera was measured in Google Earth Pro using the 3D path

measuring tool at 40 control points in the image. These distance points were then linearly interpolated in Matlab to calculate a distance to each pixel. Water ε values over a range of viewing angles were compiled (Sidran, 1981; Masuda et al., 1988; Sobrino and Cuenca, 1999). TIR temperatures were corrected assuming emissivities based on camera angle. While wind speed can also affect water surface emissivity, we did not vary emissivity as a function of changing winds for simplicity. The variation in ε as a function of wind speed was relatively small for the two field periods: 0.804 (0 m/s) to 0.853 (15 m/s) for 2015 and 0.879 (0 m/s) to 0.889 (15 m/s) for 2016 (Masuda et al., 1988). Reflected temperatures were initially assumed to equal air temperatures similar to previous time-lapse, ground-based TIR studies (Cardenas et al., 2014; Aubry-Wake et al., 2015).

An empirical correction method was developed to determine if it provides a more effective way to correct for reflected radiation when performing ground-based infrared surveys. This method used a temperature adjustment based on the average difference between kinetic and radiant stream temperatures for a subset of three control points. Control points 20, 26, and 32 were used for 2015 and control points 1, 7, and 12 were used for 2016; these control points are located approximately at the center and edges of each set of images. The uncorrected radiant temperature at each control point was calculated as the average of a cluster of 9 pixels extracted from the center of the stream. The average residual of the three control point locations at each time was used to adjust radiant stream temperatures throughout each image. Empirically adjusted radiant temperatures were compared to kinetic temperatures to assess the effectiveness of the method. Pixels along the midline of the stream were extracted to analyze longitudinal temperature patterns within the empirically corrected images (Vatland et al., 2015).

2.5 Estimation of Reflected Temperatures

In the analytical correction approach, reflected temperatures are assumed to equal air temperatures, based on the methods from previous studies that used time-lapse, ground-based TIR imagery (Cardenas et al., 2014; Aubry-Wake et al., 2015). To assess whether this was an accurate assumption, we calculated the reflected temperatures needed for TIR temperatures (T_{meas}) corrected using Eq. (1) to equal the instream temperatures measured by the iButton sensors. We estimated these reflected temperatures by assuming the radiant temperature of the stream ($T_{surface}$) equaled the kinetic temperature recorded by the iButton sensors and solving for the reflected temperatures (T_{refl}) through time using Eqs. (1) and (2), constant emissivity, and calculated atmospheric transmissivity values.

2.6 Reflected Temperature Experiment

To assess whether these estimated reflected temperatures are reasonable and the range of possible reflected radiant sky temperatures under different conditions, TIR images of 0.92×0.92 cm cardboard covered in crumpled aluminum foil were taken in the center of the Syracuse University quad (43.0376°N, 76.1340°W) away from buildings or overhanging vegetation. We recorded TIR images at different times over several days using a FLIR One camera. Aluminum foil has an emissivity of ~zero (0.03–0.07; Lillesand et al., 2015), so all radiant energy from the foil is reflected from the environment, rather than related to the temperature of the foil. Images were taken under different sky conditions (clear, scattered clouds, cloudy). Concurrent weather data were downloaded from the Syracuse University meteorological station (https://onondaga.weatherstem.com/syracuse#, 43.0382°N, 76.1334°W). The radiant temperatures of the foil pixels were then analyzed to determine the range in reflected sky temperatures. To determine the effect reflected temperatures have on TIR data accuracy, Eq. (1)

(τ=1) was used to calculate theoretical measured radiant temperatures of specific kinetic temperatures, for a range of reflected temperatures and water surface emissivity values. The relationship between the reflected temperature and water surface emissivity, along with the measured reflected sky temperatures, were used to assess the estimated reflected temperatures for the TIR images of the Quilcay Stream.

3 Results and Discussion

3.1 Uncorrected Thermal Infrared Stream Temperatures

The uncorrected TIR images do not provide accurate stream temperature data. The maximum, minimum, and mean radiant temperatures from clusters of 9 pixels at each control point show similar temperatures and so the mean was used (Fig. 5). Uncorrected radiant stream temperatures have mean absolute errors of 5.67 °C (2015) and 3.53 °C (2016), with radiant temperatures typically colder than kinetic temperatures (Fig. 5a-b). This difference between the in-stream sensors and TIR images is not due to thermal stratification, because the three stream temperature profiles never differ by>0.32 °C, and two profiles never differ by>0.12 °C. In addition, this error is too large to be from the presence of a thermal boundary layer/skin temperature, which is typically only 0.1–0.5 °C less than the bulk water temperature (Zappa and Jessup, 1998; Minnet et al., 2001; Jessup and Branch, 2008) and would be disrupted and mixed within the turbulent stream.

While TIR errors vary through time, errors at any given time are similar for all control points. Morning radiant stream temperatures rise prior to in-stream temperatures (Fig. 5a-b). In another study, TIR glacial temperatures rose prior to direct glacial surface temperatures, which was attributed to reflected radiation (Aubry-Wake et al., 2015). We believe the premature morning warming of radiant stream temperatures in the current study is due to reflected radiation

from solar heating of the cliffs and banks, which warm earlier than the stream on the valley floor (Puleo et al., 2012). Two periods during 2015 exhibit TIR temperatures out of phase with the diurnal temperature signal (Fig. 5a). Unlike the 2015 data, all 2016 TIR data follow the correct diurnal pattern. The weather data from the two field periods are very similar (Fig. 1), and so we do not believe these out of phase periods are due to changes in air temperature or relative humidity that affect the atmospheric transmissivity. Therefore, these out of phase periods may result from the more oblique 2015 viewing angle and therefore more reflection, or be due to interference from some atmospheric or surficial phenomena we failed to record. Overall, the differences between the radiant and kinetic temperatures seem mainly the result of the interaction between reflected temperatures from the surrounding environment and water surface emissivity, with the influence of reflection stronger in 2015 due to the more oblique viewing angle (Puleo et al., 2012).

3.2 Analytically Corrected TIR Stream Temperatures

TIR stream temperatures were analytically corrected for atmospheric transmissivity (τ), stream surface emissivity (ε), and reflected temperatures (T_{refl}) using Eq. (1). For this analytical correction method, emissivity was first assumed to equal 0.96 according to previous hydrologic time-lapse, ground-based TIR studies (Cardenas et al., 2011; Cardenas et al., 2014). The analytical correction resulted in mean absolute temperatures differences of 6.27 °C and 3.79 °C for 2015 and 2016 (Fig. 5c–d). Viewing angle based emissivity values of 0.80 (~75°) and 0.88 (~70°) were then used to correct the raw 2015 and 2016 TIR data (Masuda et al., 1988), producing mean absolute differences of 8.24 °C and 4.00 °C (Fig. 5c–d). Even if we had incorporated changing emissivity due to variations in wind speed at these viewing angles, daytime TIR stream temperatures, which occurred during periods of high wind, would only be

slightly warmer than when emissivity was held constant according to the viewing angle: <1.5 °C warmer (ε=0.853 with wind of 15 m/s) in 2015 and < 0.5 °C warmer (ε=0.889 with wind of 15 m/s) in 2016 (Masuda et al., 1988). Therefore, regardless of the emissivity value used, corrected stream temperatures from the analytical method were colder than the uncorrected TIR data and further from the kinetic stream temperatures. The two variables in Eq. (1) with the greatest uncertainty are emissivity and reflected temperature; the values of these variables are assumed, while the rest are either measured or calculated. Since corrected radiant temperatures remain too cold when emissivity is decreased, we conclude it is incorrect to assume the reflected temperatures equal the air temperatures.

3.3 Relationship Between Emissivity and Reflected Temperature

Since water surface emissivity varies with viewing angle (Fig. 4a; Sidran, 1981; Masuda et al., 1988; Sobrino and Cuenca, 1999), we plotted the temperature difference between a kinetic stream temperature of 7°C and the measured radiant temperature resulting from a range of emissivity values and reflected temperatures (Fig. 6b). When emissivity is ∼1, radiant temperatures measured by the camera equal kinetic temperatures, regardless of the reflected temperature. As emissivity decreases, the reflected temperature has a greater influence on the measured TIR temperature. If the reflected temperature is warmer than the stream ($T_{refl}–T_{actual} > 0$), measured radiant temperatures will exceed kinetic temperatures ($T_{meas}–T_{actual} > 0$). If the reflected temperature is colder than the stream ($T_{refl}–T_{actual} < 0$), measured radiant temperatures will be colder than kinetic temperatures ($T_{meas}–T_{actual} < 0$; Fig. 6b; Saunders, 1967). If the kinetic stream temperature is 7°C, and emissivity equals 0.8 with reflected temperatures 10°C colder than the stream, TIR camera measurements will have an error of −1.9°C, compared to the −0.4°C error when emissivity equals 0.96 (Fig. 4b; points x & y). Even from nadir views (ε=0.985),

studies measuring ocean temperatures found reflections on days with clear skies could result in uncorrected temperatures up to 0.5°C colder than reality (Zappa and Jessup, 1998). Therefore, it is important to correctly determine the emissivity and reflected temperatures to accurately correct radiant stream temperatures.

While accurate surface emissivity values and reflected temperatures are important for absolute temperature corrections, they are also important for comparing relative temperatures. This is especially important if a stream is viewed obliquely, as emissivity will vary along the stream due to viewing angle variation. If a stream has a homogenous temperature of 7°C, but is viewed obliquely, the portion of the reach viewed at a lower angle (75°) has an emissivity of 0.8 and a TIR error of −3.2°C if reflected temperatures are 17°C colder than the stream (Fig. 6b point w). Meanwhile, the portion viewed at a less oblique angle (70°) has an emissivity of 0.88 and a TIR error of −1.9°C (Fig. 6b point z). This results in a 1.3°C temperature difference along the reach due to viewing angle differences, rather than from actual temperature differences. As emissivity decreases and the reflected temperatures become increasingly different from the stream temperatures, TIR errors increase. This problem becomes more complex if the reflected temperature from the surrounding environment varies along the reach.

3.4 Estimated Reflected Temperatures of Quilcay Stream Imagery

We estimated the reflected temperature needed for the analytically corrected TIR data to equal the temperatures recorded by the in-stream sensors at our study site. By assuming $T_{surface}$ equaled the in-stream temperature, we solved for the reflected temperature through time using Planck's Law and Eq. (1) with constant emissivity and calculated atmospheric transmissivity. Calculated estimates of reflected temperatures for the Quilcay Stream data typically ranged from about −50 to 21°C for 2015 and from −60 to 18°C for 2016 (Fig. S1). For both datasets,

estimated reflected temperatures were typically warmer during the day than night, though sometimes cold reflected temperatures occurred during the day, particularly during 2015. The reflected temperatures required to generate the observed radiant temperatures are not consistent with the assumption that the reflected temperature is well represented by observed air temperatures (Cardenas et al., 2011; Cardenas et al., 2014; Aubry-Wake et al., 2015).

3.5 Observed Reflected Radiant Sky Temperatures

To assess whether these inferred reflected temperatures are reasonable, we measured reflected radiant sky temperatures under a range of conditions and times of day since the sky was the likeliest source of the cold reflected temperatures we estimated. Our observations show average reflected sky temperatures were colder than corresponding air temperatures, and minimum reflected temperatures were as cold as approximately −40°C (Fig. S2). Similarly, Chen and Zhang (1989) measured average reflected sky temperatures of −9.12°C and found sky temperatures were often>30°C cooler than ground temperatures. Other studies also cite radiant reflected clear sky temperatures of −18°C (Zappa and Jessup, 1998) and −20 to −50°C depending on location (Jessup and Branch, 2008). Our measurements, along with results from previous studies, indicate the reflected sky temperatures are generally colder under clear skies and warmer under cloudy conditions (Fig. 7; Fig. S3; Saunders, 1967; Zappa and Jessup, 1998). Under scattered clouds, reflected sky temperatures exhibit more variability due to patches of sky with and without clouds (Fig. 7). Different cloud types located at different heights will emit different radiative temperatures (Donlon et al., 2008).

Since the sky is a main source of the reflected temperature emitted from the environment at the field site in this study, the range in TIR temperatures that the sky exhibits influences the measurement of radiant stream temperatures. For the Quilcay site, reflected temperatures could

be even colder than those measured in Syracuse since reflected sky temperatures are colder at higher elevation due to the thinner, drier overlying atmosphere (Smith et al., 1996; Minnet et al., 2001). Other objects in the surrounding environment also produce reflected temperature errors, especially during the day under clear sky conditions (Marruedo et al., 2018). Other sources of reflections are the cliffs and gravel banks, which can sometimes emit the highest radiative temperatures (Tonolla et al., 2010); during the day, temperatures of the gravel bar and dry areas of the meadow can exceed 30°C based on iButton data, while at night ground temperatures can drop below 5°C. Therefore, the range of reflected temperatures we estimated for the Quilcay site are reasonable and typically not equal to air temperatures. Furthermore, when surface roughness is low, such as under low winds and calm water surfaces, reflections are pronounced. Such reflections are less discernable under high winds and rough water surfaces (Zappa and Jessup, 1998). This could also explain why the measured nighttime stream temperatures, which occur when the wind speed drops, exhibit TIR errors that are colder than daytime errors. Estimated daytime reflected temperatures may also be warmer than estimated nighttime reflected temperatures due to the mix of sky reflections with reflections from the cliffs and surrounding land surface, which emit warm radiative temperatures during the day and cooler radiative temperatures at night. While the sky was the main source of reflected radiation in our study, other sources could be more important at sites with different characteristics. For example, on a stream mostly blocked by trees with minimal view of the sky, the reflected temperature would mainly come from the surrounding trees and vegetation, often resulting in warmer reflected temperatures than those from the sky.

3.6 Empirically Corrected TIR Temperatures

Since the analytical correction method can be inaccurate due to errors estimating water surface emissivity and reflected temperature, we also explored an empirical correction method. Three control points were used to determine the offset value of the stream temperatures from each TIR image through time. The remaining control points (n=4 in 2015 and n=7 in 2016) were then used to assess the resulting error of empirically corrected stream temperatures. The mean absolute temperature differences of the control points not used to determine the offset correction were 0.30°C and 0.16°C for 2015 and 2016 (Fig. 5e–f). The empirical method reduced radiant stream temperature errors, but certain time periods are more error prone, particularly mornings in 2015 (Fig. 5g–h). This empirical correction likely worked better in 2016 than in 2015 because the 2016 reach was shorter and so had less spatial variation in the reflected temperature and viewing angle, making errors more consistent at any time. Overall, the empirical correction improved ground-based radiant temperatures more than the analytical correction similar to findings from aerial studies that used calibrations based on direct water temperature measurements (Wawrzyniak et al., 2013).

While the empirical correction method is more effective at reducing the TIR stream temperature error, limitations remain. The empirical method does well correcting image scale reflections (e.g. sky and cliff reflections) but it cannot correct smaller scale reflections including those from patchy cloud cover, objects such as trees and tall vegetation near the stream edge, or rocky stream banks that emit large amount of radiant energy that interfere with the stream's thermal signal (Puleo et al., 2012). In places where an insufficient number of pixels span the stream, errors occur due to mixed pixels or thermal contamination from bank reflections. Previous work shows at least 10 pixels should span the stream to obtain uncontaminated

temperatures (Torgersen et al., 2001). Here, we limited our results to locations where there are 8 or more pixels spanning the width of the stream. A final drawback of the empirical method is that the offset correction is unique to each data set and TIR image, and the correction factor changes through time due to factors such as weather conditions and site characteristics. Therefore, direct temperature measurements must be made to create the offset correction applied to each TIR image. Additionally, if you are interested in measuring the TIR temperature of multiple surface types (e.g. water, rock) in an image you must apply a different correction factor to each surface type because their errors will be different due to their differing emissivities.

3.7 TIR Stream Temperature Patterns

Kinetic temperature measurements show that Quilcay Stream temperatures are fairly uniform spatially, despite the inflow of small tributaries (Movie S1; Movie S2). However, corrected radiant temperatures sometimes show warmer and cooler spots that are unsupported by the in-stream measurements. Particularly problematic times occur when temperatures are rising or falling, especially near where the stream narrows (Movie S1 upstream of control point 27; Movie S2 control points 2–5). Stream segments not spanned by at least 8 pixels have larger temperature anomalies during these periods likely due to thermal contamination from the stream bank. Additionally, the reach between control points 10 and 11 sometimes shows artificially fast warming in the mornings due to thermal contamination from the adjacent gravel banks (Movie S2).

While longitudinal stream temperatures typically show minor noise (±0.5 °C), some periods during 2015 exhibit increased noise (e.g. July 22 from 3:00–6:00, July 24 at 14:50) due to reflections caused by the oblique viewing angle (Puleo et al., 2012). Overall, false temperature

anomalies are more prevalent for 2015 than 2016, likely due to the more oblique viewing angle that decreases water surface emissivity and due to fewer buffer pixels between the stream banks and stream midline, allowing for more contamination from bank reflections.

3.8 Comparison to Previous Time-Lapse, Ground-Based Studies

Few studies have used time-lapse, ground-based TIR imagery to obtain kinetic stream surface temperatures, but these studies have demonstrated similar TIR temperature correction issues. Data from a time-lapse, ground-based TIR study assessing glacial surface temperatures also exhibits periods when the radiant and kinetic temperatures are out of phase, and radiant temperatures warm in the morning prior to warming of kinetic temperatures (Aubry-Wake et al., 2015). TIR data from the study were adjusted using a 5.96°C offset correction after the analytical correction was applied (ε_{snow}=0.98, ε_{ice}=0.97, T_{refl}=air temperature), indicating the inputs used in the analytical method did not account for reflected temperatures that were colder than air temperatures and lower emissivity due to variations in the glacial surface.

Another study used time-lapse, ground-based TIR imagery to analyze radiant stream temperature data over ~24 hours (Cardenas et al., 2014). Similarly, morning radiant temperatures rose more quickly than kinetic in-stream temperatures. The study has a viewing angle comparable to our 2015 images, with an oblique view that varies along the stream length and prevents numerous pixels from spanning the stream. This could result in erroneous stream temperatures due to thermal contamination and assumptions about emissivity and reflected temperature values (ε=0.96 and T_{refl}=air temperatures) since the viewing angle varies (~65–85°) in these images, with more oblique views of the upstream and downstream portions of the reach. Since the study only has one control point which is obscured from the camera's view, it is uncertain whether observed upstream and downstream warming was due to warm inflows, as the

authors concluded, or due to lower upstream and downstream water surface emissivity values enabling a greater proportion of the surrounding land temperatures (which were always warmer than the stream) to reflect off the stream, thereby making these regions appear warmer.

3.9 Proposed Methodological Changes

Previous time-lapse, ground-based TIR studies have made incorrect generalizing assumptions about stream surface emissivity and reflected temperature values (ε=0.96 and T_{refl}=air temperatures), leading to applications of the analytical method that produced TIR temperatures that may be incorrect, or correct for the wrong reasons. For example, errors due to incorrectly high emissivity values at oblique viewing angles may be hidden if reflected temperatures are close in value to the temperature of the target object (Zappa and Jessup, 1998). Experimental design should also be carefully considered to achieve high viewing angles and avoid designs where the viewing angle of the stream varies across the image altering the emissivity. If the hydrologic community continues to use time-lapse, ground-based TIR imagery for extracting accurate stream temperatures (e.g. for comparison to stream temperature energy balance model output), then methodological changes must be considered.

One possible methodological change is implementation of the empirical approach instead of the analytical correction approach, such that results do not rely on accurate values of emissivity and reflected temperature, especially in environments where these variables are difficult to obtain. Another possibility is to draw on methodological best practices for TIR sensing of sea surface temperature (SST). Prior work in this area has demonstrated that accurate emissivity and reflected temperature values must be known to make accurate temperature corrections using the analytical method (e.g. Smith et al., 1996; Zappa and Jessup, 1998; Minnet et al., 2001; Jessup and Branch, 2008; Donlon et al., 2008). To measure reflected temperatures,

prior investigators have used either multiple radiometers positioned at the same viewing angle to the sea and sky (Smith et al., 1996; Zappa and Jessup, 1998; Minnet et al., 2001; Jessup and Branch, 2008) or dual port radiometers – one radiometer with a rotating mirror that alternately directs sky and surface radiance into the detector (Donlon et al., 2008). While the methods for TIR sensing of SST still have limitations (e.g. time of sky and sea measurement must be the same, rough water surfaces with various orientations can receive different reflections, greater spatial variability in reflected temperatures under scattered cloud conditions; Jessup and Branch, 2008; Donlon et al., 2008), they provide a methodology that may be an improvement from the incorrect assumptions often made when using the analytical approach for sensing stream and surface water temperatures.

4 Conclusions

Improvements in the accuracy and cost-effectiveness of TIR temperature sensing present an opportunity to obtain high-resolution temperature records of streams and surface water through space and time. Such data can be used to map thermal heterogeneities in streams and to inform model simulations of stream temperatures. However, accurate TIR sensing of stream and surface water temperatures remains a challenge due to the complexities surrounding correction of time-lapse, ground-based TIR imagery. Analytical corrections of TIR temperatures that rely on assumed values of water emissivity and reflected temperatures are often ineffective because it is difficult to obtain a near zenith view of a stream from the ground. This leads to uncertainty in the emissivity values along stream reaches due to the non-linear relationship between emissivity and viewing angle. As emissivity decreases, correct quantification of reflected temperatures becomes increasingly important and cannot be assumed to equal air temperature. However, it is

difficult to measure reflected temperatures along an entire stream reach as they can vary both spatially and temporally, though measurement of reflected temperatures could be improved by using multiple or dual port radiometers. Empirical corrections based on the observed offset between in situ sensors and TIR (radiant) temperatures are partially effective at correcting radiant stream temperatures as this method corrects image scale reflections without needing the emissivity and reflected temperature values to be known. However, neither method corrects for small scale reflections that can occur, such as those from localized gravel bars or overhanging vegetation. Both methods have greater error when the viewing angle varies along the stream reach and when too few pixels span the stream width, allowing thermal contamination from the banks. Due to uncertainty in reflected temperatures and emissivity values, it may be difficult to accurately measure absolute stream temperatures and temperature differences along a reach using ground-based TIR sensing. Therefore, while ground-based TIR remains informative for identifying locations of relative stream temperature difference, caution should be taken when drawing conclusions about environmental processes from absolute TIR temperature data.

References

Anderson, J.M., Duck, R.W., McManus, J., 1995. Thermal radiometry: a rapid means of determining surface water temperature variations in lakes and reservoirs. *J. Hydrol*. 173 (1-4), 131–144. https://doi.org/10.1016/0022-1694(95)02714-Z.

Atwell, B.H., MacDonald, R.B., Bartolucci, L.A., 1971. Thermal mapping of streams from airborne radiometric scanning. *Water Resour. Bull*. 7 (2), 228–243.

Aubry-Wake, C., Baraer, M., McKenzie, J.M., Mark, B.G., Wigmore, O., Hellström, R.Å., Lautz, L., Somers, L., 2015. Measuring glacier surface temperatures with ground based

thermal infrared imaging. *Geophys. Res. Lett*. 42, 8489–8497. https://doi.org/10.1002/2015GL065321.

Aubry-Wake, C., Zéphir, D., Baraer, M., McKenzie, J.M., Mark, B.G., 2018. Importance of longwave emissions from adjacent terrain on patterns of tropical glacier melt and recession. *J. Glaciol*. 64 (243), 49–60. https://doi.org/10.1017/jog.2017.85.

[dataset] Baker, E.A., Lautz, L.K., McKenzie, J.M., 2019. Improving the accuracy of time lapse thermal infrared imaging for hydrologic applications. *HydroShare*, http://www.hydroshare.org/resource/6b12cc90422b41619f686f2d3bdade37.

Banks, W.S.L., Paylor, R.L., Hughes, W.B., 1996. Using thermal-infrared imagery to delineate ground-water discharge. *Ground Water* 34 (3), 434–443.

Bingham, Q.G., Neilson, B.T., Neale, C.M.U., Cardenas, M.B., 2012. Application of high resolution, remotely sensed data for transient storage modeling parameter estimation. *Water Resourc. Res*. 48 (8), 1–15. https://doi.org/10.1029/2011WR011594.

Buettner, K.J.K., Kern, C.D., 1965. The determination of infrared emissivities of terrestrial surfaces. *J. Geophys. Res*. 70 (6), 1329–1337. https://doi.org/10.1029/JZ070i006p01329.

Cardenas, M.B., Neale, C.M.U., Jaworowski, C., Heasler, H., 2011. High-resolution mapping of river-hydrothermal water mixing: Yellowstone National Park. *Int. J. Remote Sens*. 32 (10), 2765–2777. https://doi.org/10.1080/01431161003743215.

Cardenas, M.B., Doering, M., Rivas, D.S., Galdeano, C., Neilson, B.T., Robinson, C.T., 2014. Analysis of the temperature dynamics of a proglacial river using time-lapse thermal imaging and energy balance modeling. *J. Hydrol*. 519, 1963–1973. https://doi.org/10.1016/j.jhydrol.2014.09.079.

Chen, J., Zhang, R., 1989. Studies on the measurements of crop emissivity and sky temperature.

Agric. For. Meteorol. 49, 23–34.

Donlon, C., Robinson, I.S., Reynolds, M., Wimmer, W., Fisher, G., Edwards, R., Nightingale, T.J., 2008. An infrared sea surface temperature autonomous radiometer (ISAR) for deployment aboard volunteer observing ships (VOS). *J. Atmospheric Oceanic Technol*. 25 (1), 93–113. https://doi.org/10.1175/2007JTECHO505.1.

Dugdale, S.J., 2016. A practitioner's guide to thermal infrared remote sensing of rivers and streams: recent advances, precautions and considerations. *Wiley Interdiscip. Rev.: WIREs Water* 3, 251–268. https://doi.org/10.1002/wat2.1135.

Eschbach, D., Piasny, G., Schmitt, L., Pfister, L., Grussenmeyer, P., Koehl, M., Skupinski, G., Serradj, A., 2016. Thermal-infrared remote sensing of surface-water/groundwater exchanges in a restored anastomosing channel (Upper Rhine River, France). *Hydrolog. Processes* 31, 1113–1124. https://doi.org/10.1002/hyp.11100.

Fricke, K., Baschek, B., 2015. Temperature monitoring along the Rhine River based on airborne thermal infrared remote sensing: estimation of in situ water temperatures and inflow detection compared to artificial satellite data. *J. Appl. Remote Sens*. 9 (1), 95067. https://doi.org/10.1117/1.JRS.9.095067.

Gillespie, A., Rokugawa, S., Matsunaga, T., Steven Cothern, J., Hook, S., Kahle, A.B., 1998. A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images. IEEE Trans. *Geosci. Remote Sens*. 36 (4), 1113–1126. https://doi.org/10.1109/36.700995.

Hare, D.K., Briggs, M.A., Rosenberry, D.O., Boutt, D.F., Lane, J.W., 2015. A comparison of

thermal infrared to fiber-optic distributed temperature sensing for evaluation of groundwater discharge to surface water. *J. Hydrol*. 530, 153–166. https://doi.org/10.1016/j.jhydrol.2015.09.059.

Handcock, R.N., Gillespie, A.R., Cherkauer, K.A., Kay, J.E., Burges, S.J., Kampf, S.K., 2006. Accuracy and uncertainty of thermal-infrared remote sensing of stream temperatures at multiple spatial scales. *Remote Sens. Environ*. 100 (4), 427–440. https://doi.org/10.1016/j.rse.2005.07.007.

Handcock, R.N., Torgersen, C.E., Cherkauer, K.A., Gillespie, A.R., Tockner, K., Faux, R.N., Tan, J., 2012. Thermal infrared remote sensing of water temperature in riverine landscapes. *Fluvial Remote Sens. Sci. Manage*. 85–113. https://doi.org/10.1002/9781119940791.ch5.

Jessup, A.T., Branch, R., 2008. Integrated ocean skin and bulk temperature measurements using the calibrated infrared in situ measurement system (CIRIMS) and through-hull ports. *J. Atmospheric Oceanic Technol*. 25 (4), 579–597. https://doi.org/10.1175/2007JTECHO479.1.

Kay, J.E., Kampf, S.K., Handcock, R.N., Cherkauer, K.A., Gillespie, A.R., Burges, S.J., 2005. Accuracy of lake and stream temperatures estimated from thermal infrared images. *J. Am. Water Resourc. Assoc*. 41 (5), 1161–1175. https://doi.org/10.1111/j.1752-1688.2005.tb03791.x.

Lillesand, T., Kiefer, R.W., Chipman, J., 2015. Remote Sensing and Image Interpretation, 6th ed. John Wiley & Sons, New York, NY.

Masuda, K., Takashima, T., Takayama, Y., 1988. Emissivity of pure and sea waters for the model sea surface in the infrared window regions. *Remote Sens. Environ*. 24, 313–329.

Marruedo Arricibita, A.I., Dugdale, S.J., Krause, S., Hannah, D.M., Lewandowski, J., 2018. Thermal infrared imaging for the detection of relatively warm lacustrine ground- water discharge at the surface of freshwater bodies. *J. Hydrol*. 562, 281–289. https://doi.org/10.1016/j.jhydrol.2018.05.004.

Minnett, P.J., Knuteson, R.O., Best, F.A., Osborne, B.J., Hanafin, J.A., Brown, O.B., 2001. The marine-atmospheric emitted radiance interferometer: A high-accuracy, seagoing infrared spectroradiometer. *J. Atmospheric Oceanic Technol* 18 (6), 994–1013. https://doi.org/10.1175/1520-0426(2001)018<0994:TMAERI>2.0.CO;2.

Mundy, E., Gleeson, T., Roberts, M., Baraer, M., McKenzie, J.M., 2017. Thermal imagery of groundwater seeps: possibilities and limitations. *Groundwater* 55 (2), 160–170. https://doi.org/10.1111/gwat.12451.

Puleo, J.A., McKenna, T.E., Holland, K.T., Calantoni, J., 2012. Quantifying riverine surface currents from time sequences of thermal infrared imagery. *Water Resourc. Res*. 48 (1), 1–21. https://doi.org/10.1029/2011WR010770.

Salisbury, J.W., D'Aria, D.M., 1992. Emissivity of terrestrial materials in the 8–14 microns atmospheric window. *Remote Sens. Environ*. 47 (3), 345–361. https://doi.org/10.1016/0034-4257(94)90102-3.

Saunders, P.M., 1967. Aerial measurement of sea surface temperature in the infrared. *J. Geophys. Res*. 72 (16), 4109–4117.

Schuetz, T., Weiler, M., 2011. Quantification of localized groundwater inflow into streams using ground-based infrared thermography. *Geophys. Res. Lett*. 38 (3), 1–6. https://doi.org/10.1029/2010GL046198.

Sidran, M., 1981. Broadband reflectance and emissivity of specular and rough water surfaces. *Appl. Opt*. 20 (18), 3176. https://doi.org/10.1364/AO.20.003176.

Smith, W.L., Knuteson, R.O., Revercomb, H.E., Feltz, W., Howell, H.B., Menzel, W.P., Nalli, N.R., Brown, O., Brown, J., Minnett, P., McKeown, W., 1996. Observations of the infrared radiative properties of the ocean — Implications for the measurement of sea surface temperature via satellite remote sensing. *Bull. Am. Meteorol. Soc*. 77 (1), 41–51.

Sobrino, A., Cuenca, J., 1999. Angular variation of thermal infrared emissivity for some natural surfaces from experimental measurements. *Appl. Opt*. 38 (18), 3931–3936.

Tonolla, D., Acuña, V., Uehlinger, U., Frank, T., Tockner, K., 2010. Thermal heterogeneity in river floodplains. *Ecosystems* 13(5),727–740. https://doi.org/10.1007/s10021-010-9350-5.

Torgersen, C.E., Faux, R.N., McIntosh, B.A., Poage, N.J., Norton, D.J., 2001. Airborne thermal remote sensing for water temperature assessment in rivers and streams. *Remote Sens. Environ*. 76 (3), 386–398. https://doi.org/10.1016/S0034-4257(01)00186-9.

Vatland, S.J., Gresswell, R.E., Poole, G.C., 2015. Quantifying stream thermal regimes at multiple scales: Combining thermal infrared imagery and stationary stream temperature data in a novel modeling framework. *Water Resourc. Res*. 51 (1), 31–46. https://doi.org/10.1002/2014WR015588.

Wawrzyniak, V., Piegay, H., Allemand, P., Grandjean, P., 2013. Prediction of water temperature heterogeneity of braided rivers using very high resolution thermal infrared (TIR) images. *Int. J. Remote Sens*. 34 (13), 4812–4831. https://doi.org/10.1080/01431161.2013.782113.

Zappa, C.J., Jessup, A.T., 1998. Skin layer recovery of free-surface wakes: Relationship to surface renewal and dependence on heat flux and background turbulence. *J. Geophys. Res*. 103, 711–722.

Tables

Table 1. Measurement periods during the two field seasons in July 2015 and August 2016.

Equipment	Measurement period	Interval (min)
Jenoptik VarioCam TIR Camera	July 20, 2015 at 16:00 to July 25, 2015 at 13:40 August 5, 2016 at 12:20 to August 9, 2016 at 6:30 (Data gap on August 7, 2016 from 2:30 to 17:30)	10
Thermochron iButtons	July 20, 2015 at 15:00 to July 26, 2015 at 8:10 August 5, 2016 at 15:40 to August 9, 2016 8:30	5
Vantage Pro2 Meteorological Station	July 20, 2015 at 13:00 to July 26, 2015 at 8:10 August 4, 2016 at 18:00 to August 9, 2016 at 8:50	10

Table S1. Technical specifications of thermal infrared (TIR) camera and temperature sensors.

TIR cameras only measure the radiant temperatures of the top 0.1 mm of a surface.

Equipment	Variable	Manufacturer	Model	Accuracy	Range
TIR Camera Peru	TIR Temp	Jenoptik	VarioCam HD 768 x 1024 pixels (uncooled microbolometer)	±1.5°C	7.5-14 μm
TIR Camera Syracuse	TIR Temp	FLIR	FLIR One Gen 3 – IOS 80 x 60 pixels	±3.0°C	8-14 μm
In-stream iButtons	Temp	Maxim	DS1922L (Thermochron iButtons)	±0.5°C (-10-65°C); 0.0625°C resolution	-40-85°C

Figures

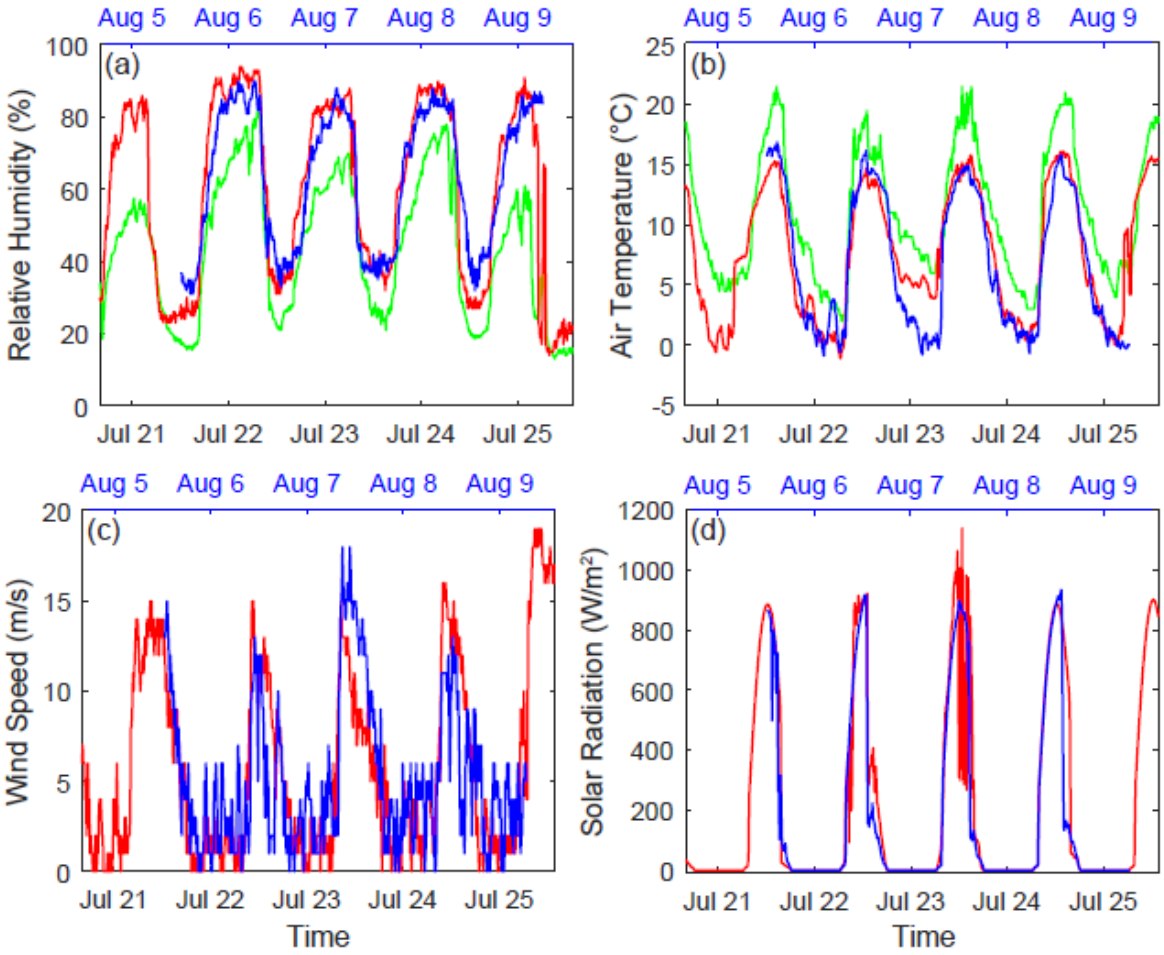

Figure 1. (a) Relative humidity, (b) air temperature, (c) wind speed, and (d) solar radiation data from Quilcayhuanca Valley during both periods of TIR image acquisition. Data in green were collected at the camera elevation during 2015, data in red were collected at the stream elevation during 2015, and data in blue were collected at the stream elevation during 2016.

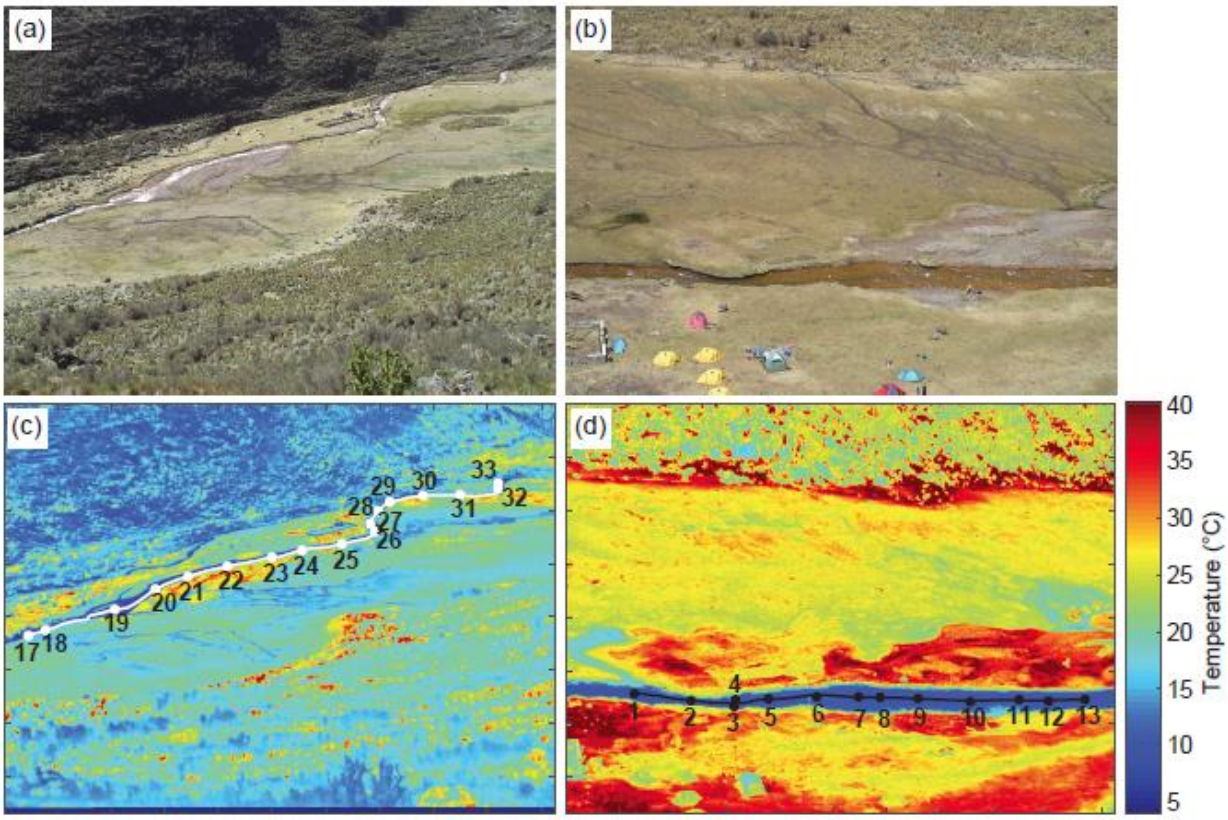

Figure 2. Visual and TIR images of the study site from (a, c) 2015 and (b, d) 2016. The control point locations are overlain on the TIR images (c-d).

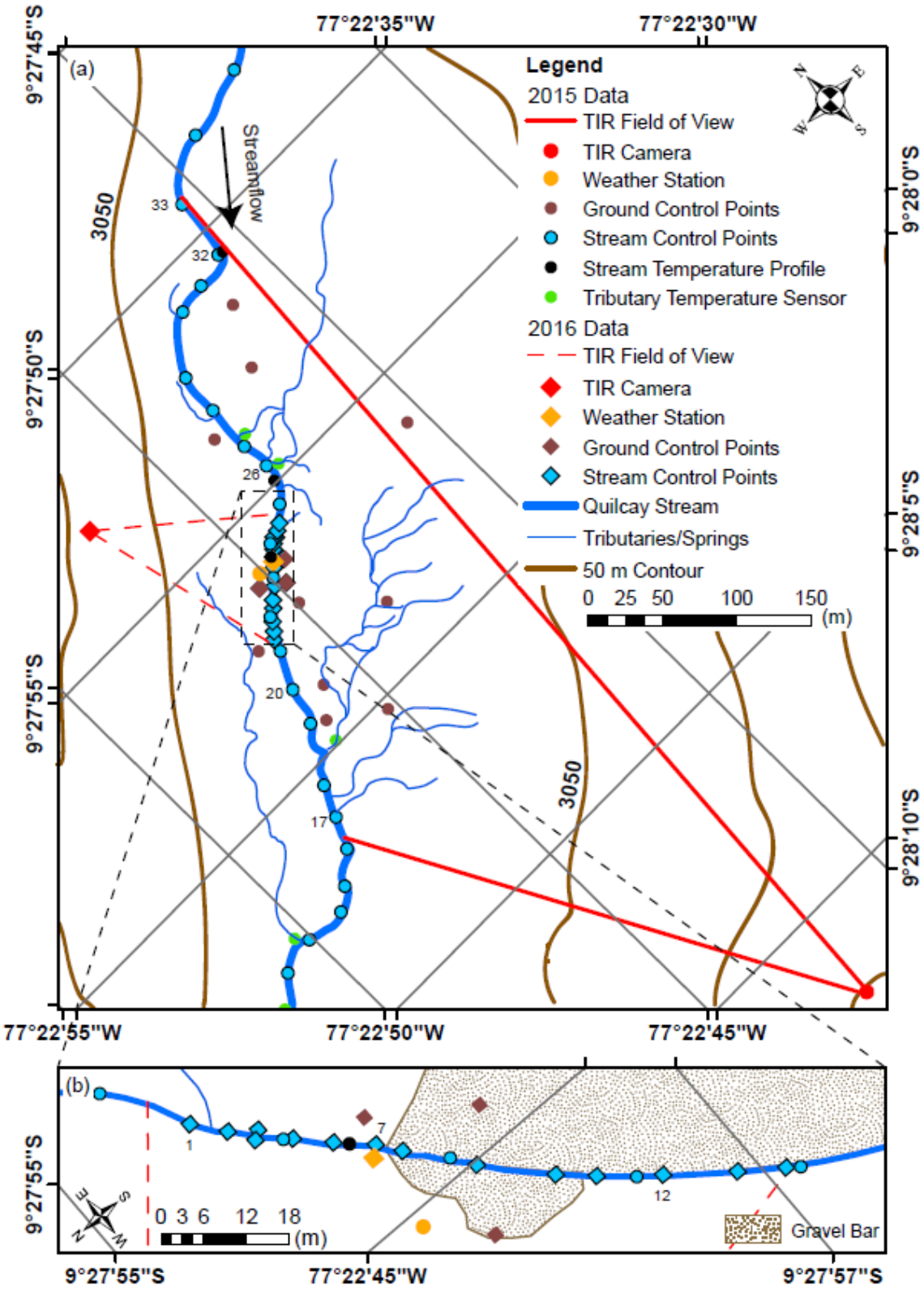

Figure 3. (a) Map of the field site depicting the TIR camera, weather station, and control point sensor locations for both field periods. (b) Magnified view of the 2016 control point locations.

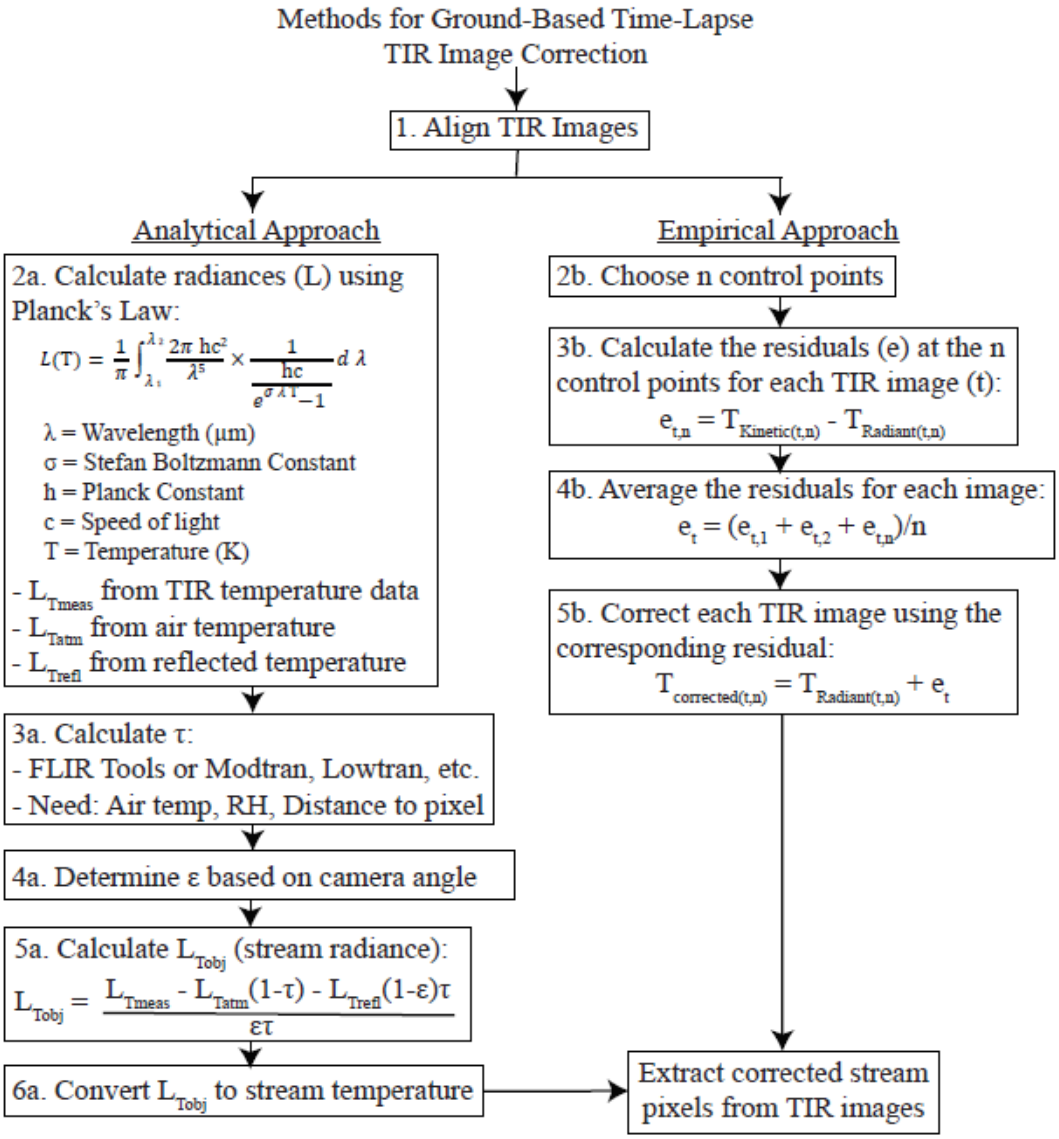

Figure 4. Flow chart of steps used to correct the ground-based, time-lapse TIR image data.

Atmospheric transmissivity values were extracted from FLIR Tools for a range of air temperature, relative humidity, and distance combinations. Air temperature and relative humidity were obtained from the weather station. The distance to the stream surface from the TIR camera was measured at 40 locations in the image using Google Earth Pro and linearly interpolated to calculate a distance to each pixel.

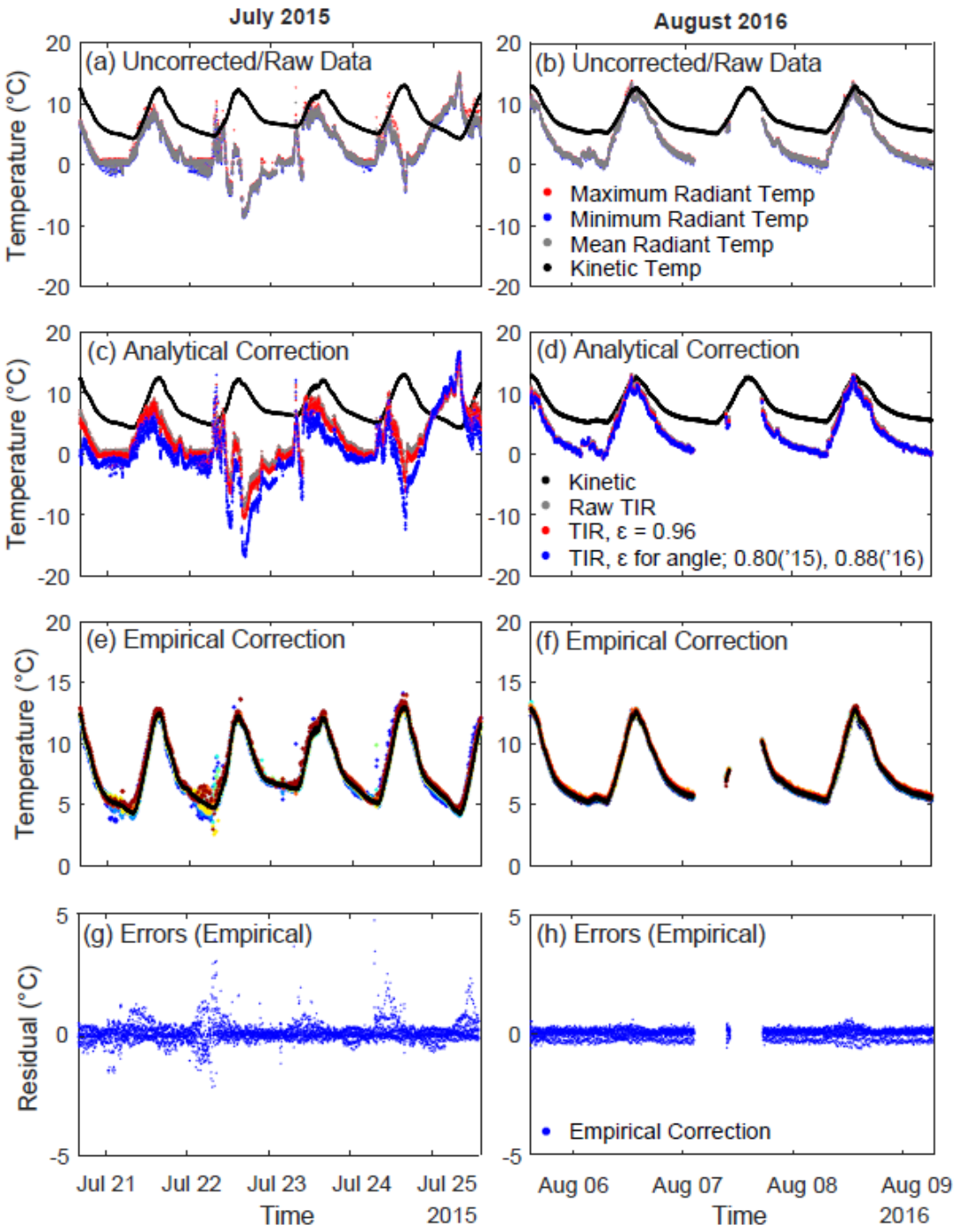

Figure 5. (a-b) Kinetic and uncorrected radiant stream temperatures at each control point. (c-d) Radiant stream temperatures after the analytical correction, assuming the reflected temperature equals air temperature. (e-f) Kinetic and radiant stream temperatures after the empirical correction. (g-h) Average errors of the empirically corrected TIR stream temperatures.

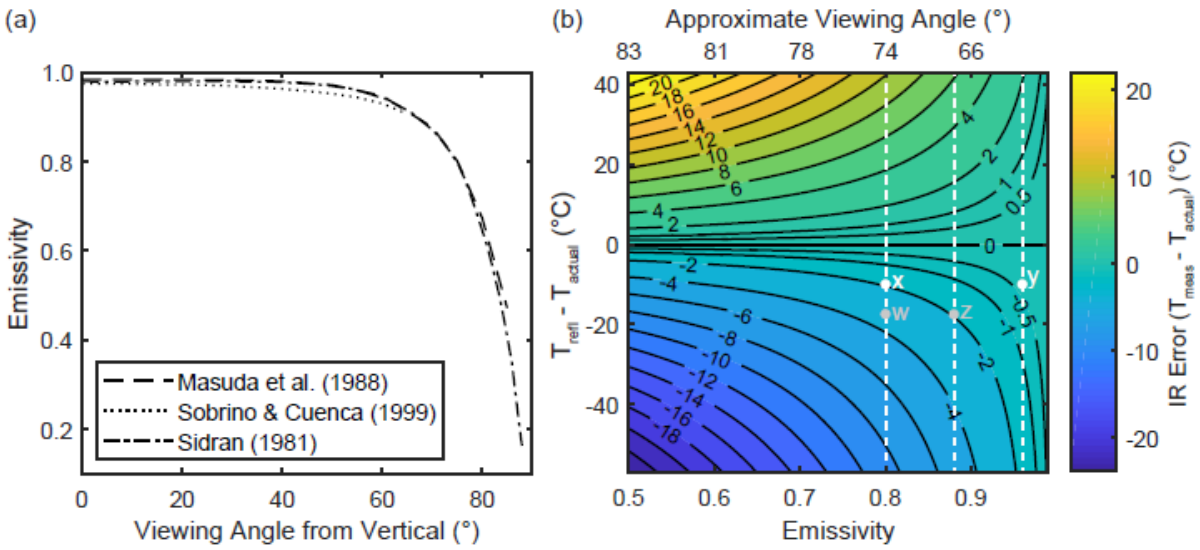

Figure 6. (a) Water surface emissivity decreases as viewing angle obliquity increases. (b) Theoretical measured radiant temperature error (T_{meas}-T_{actual}) as emissivity and the difference between the reflected temperature (T_{refl}) and stream temperature (T_{actual}) vary. The dashed lines highlight the discussed viewing angles.

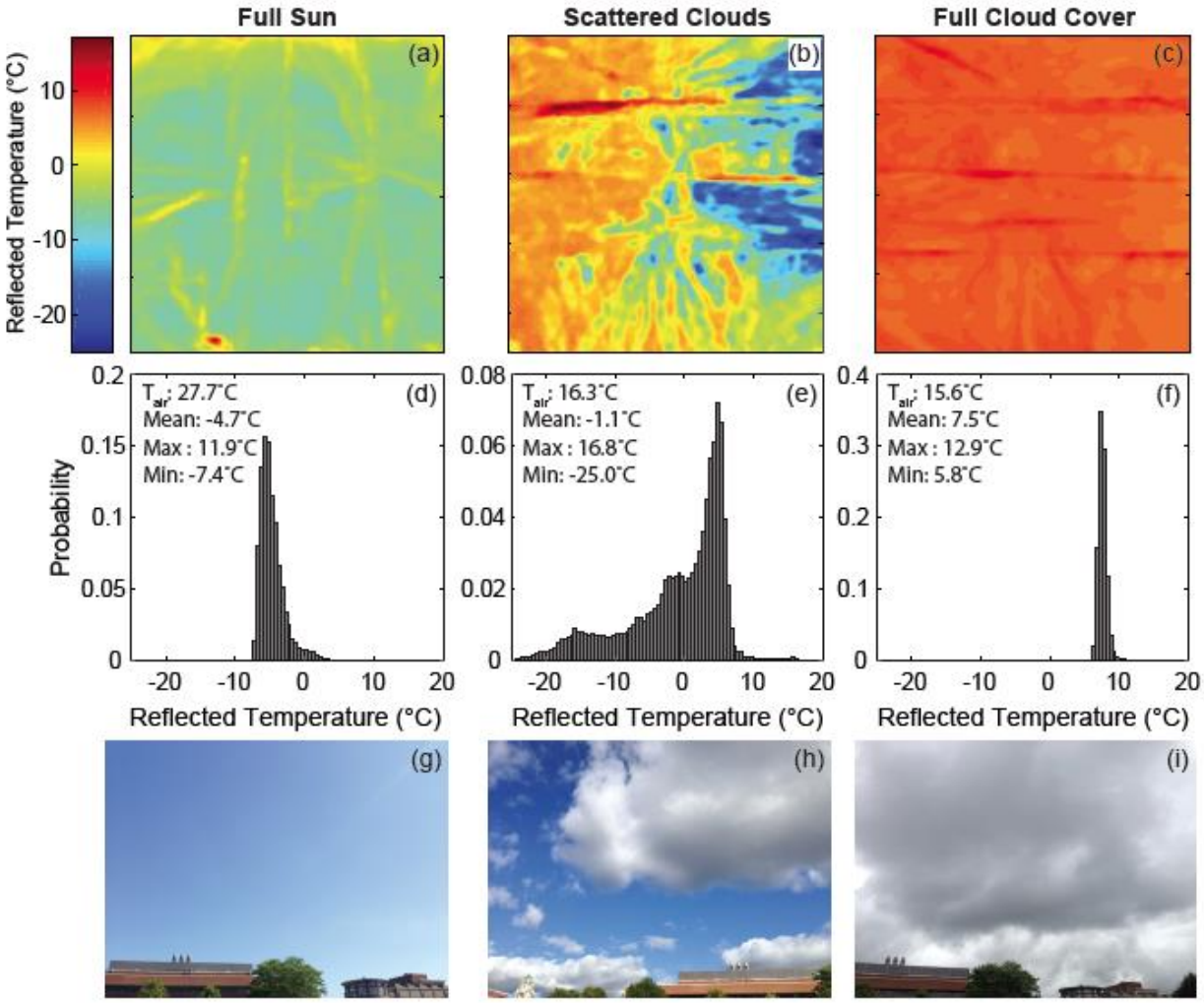

Figure 7. (a-b) TIR images of reflected sky temperatures from the sheet of foil under different conditions. (d-f) Histograms of the reflected infrared sky temperatures corresponding to the foil images. (g-i) Photographs of the corresponding sky conditions.

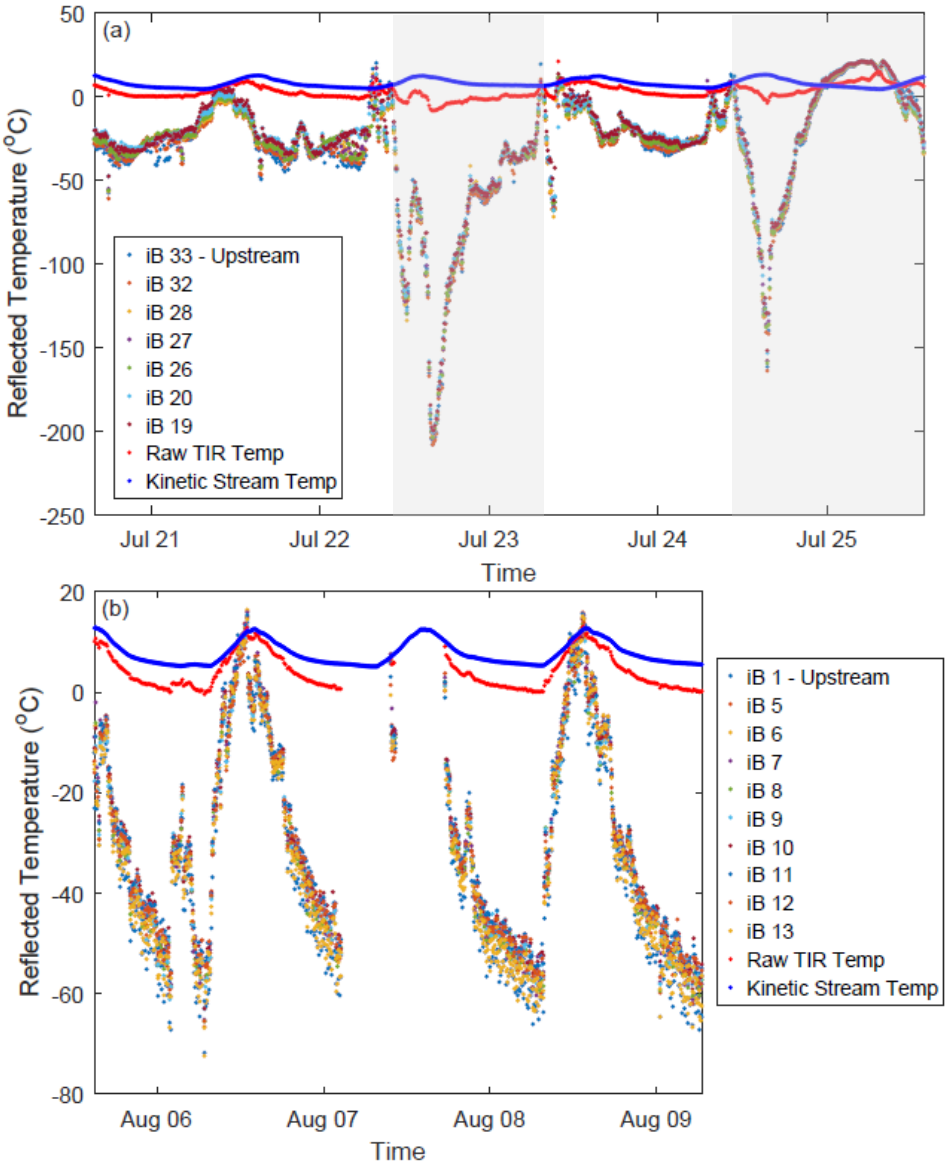

Figure S1. Estimated reflected temperatures for the (a) 2015 and (b) 2016 TIR datasets calculated by solving Eq. 1 for L_{Trefl}, assuming that $L_{Tsurface}$ is equal to the temperature measured by the in-stream temperature sensors. These are the reflected temperatures needed for the measured TIR stream temperatures (Raw TIR Temp) to match the kinetic (in-stream) temperatures when the analytical correction method is applied. The viewing angles were ~75° for 2015 and ~70° for 2016, resulting in emissivities of 0.81 and 0.88, respectively. Gray periods may have experienced some atmospheric phenomena unmeasured by the meteorological station.

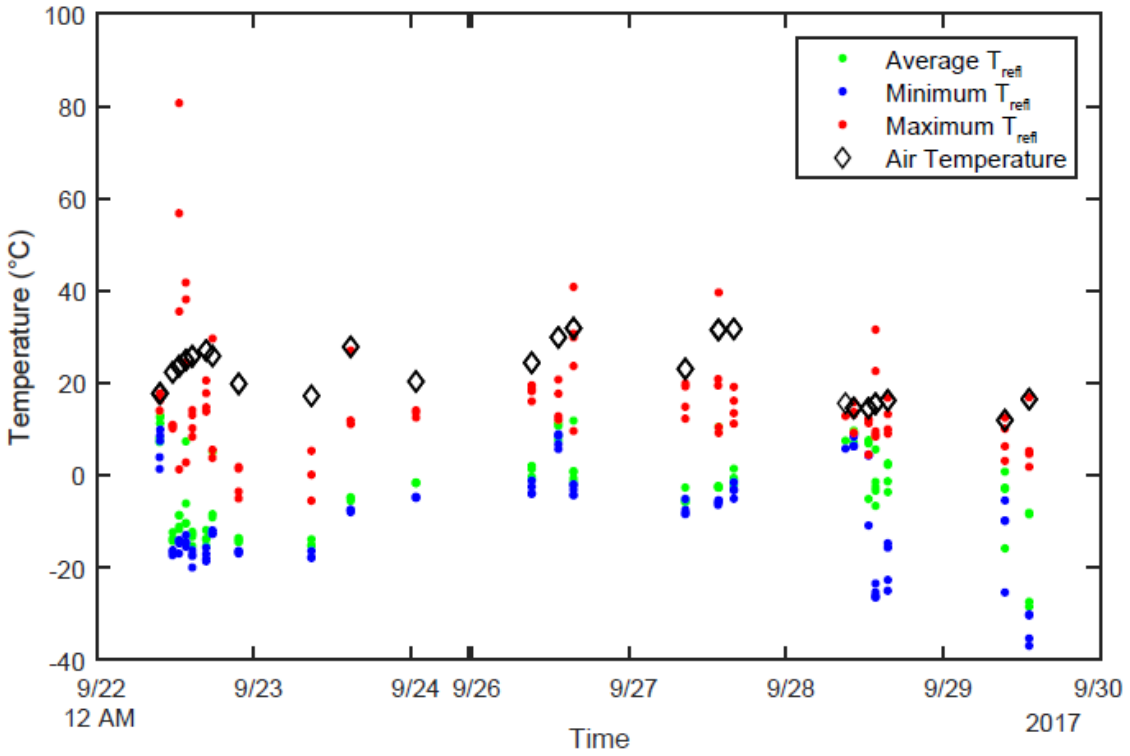

Figure S2. Measured reflected sky temperatures and air temperature data from Syracuse, NY during September 2017. Multiple TIR images (3-4) were taken of the foil reflected sky temperatures at each time. Average T_{refl} was always colder than the air temperature. Averages were taken from the whole sheet of foil, except the edges. Hot sun reflections also occur.

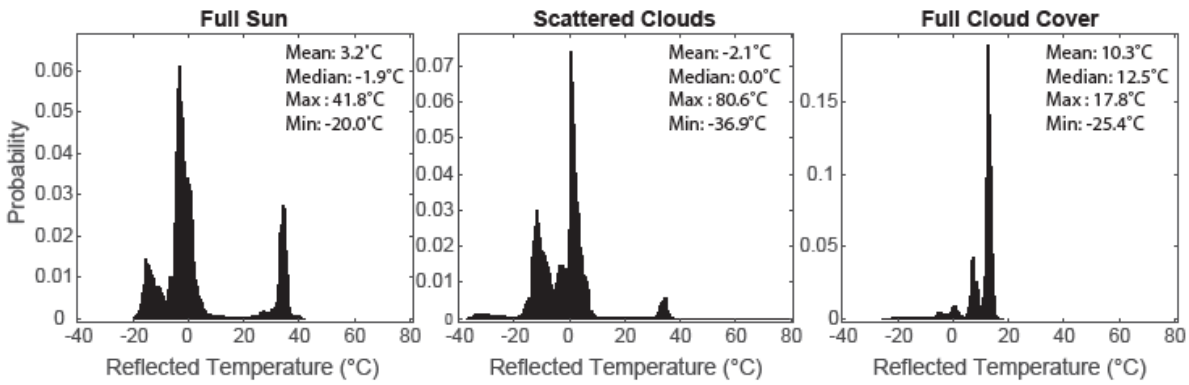

Figure S3. Histograms of reflected infrared sky temperatures from foil at 24 different times (Figure S4). Multiple images were taken at each time for a total of 92 images. TIR images of 92 x 92 cm crumpled aluminum foil covered cardboard were taken in the center of the Syracuse University quad away from buildings or vegetation (43.0376°N, 76.1340°W). High temperatures (>30°C and up to 80°C) are attributed to direct reflections from the sun; these high temperature sun reflections do not occur under full cloud cover. Sky reflections under full sun reflections are generally colder than reflections under full cloud conditions, especially when sun reflections are ignored. Under scattered cloud conditions, reflected sky temperatures are more variable and fewer hot sun reflections occur due to shielding from the clouds. Full sun was defined as less than 20% of the sky contains clouds while full cloud cover was defined as more than 80% of the sky contains clouds.

Movies

Movie S1: https://ars.els-cdn.com/content/image/1-s2.0-S0022169419301246-mmc3.mp4

Movie S1. Stream temperatures extracted from the center of the stream from the 2015 field season. Only locations where the stream was at least 8 pixels wide were used (red). Stream temperatures were corrected using the empirical correction method (an offset correction calculated from the average residuals of 3 control points at each point in time). The light gray dots depict the corrected stream temperatures at every distance down the reach. The red dots depict the corrected stream temperatures at locations downstream where a sufficient number of pixels spanned the width of the stream to extract reliable temperatures. The light blue dotted lines indicate the locations where tributaries or springs flow into the main stream.

Movie S2: https://ars.els-cdn.com/content/image/1-s2.0-S0022169419301246-mmc4.mp4

Movie S2. Empirically corrected stream temperatures extracted from the center of the stream from the 2016 field season. Only locations where the stream was at least 12 pixels wide were used (red). The temperature anomaly that occurs at ~3310 m downstream is caused by the location of the weather station. The light gray dots depict the corrected stream temperatures at every distance down the reach. The red dots depict the corrected stream temperatures at locations downstream where a sufficient number of pixels spanned the width of the stream to extract reliable temperatures. The light blue dotted line indicates the location that a spring flows into the main stream.

Chapter 3

Evaluating Groundwater Residence Time and Contributions to Streamflow in a Proglacial Alpine Catchment

Abstract

Dry season stream flow in the Peruvian Andes is predominantly sourced from a combination of glacial meltwater and groundwater, as only 20% of annual precipitation occurs from May to October. Large and small-scale agriculture, hydroelectric power, industry, and local people throughout the region depend on stream flow in the Rio Santa and its tributaries, which originates in proglacial alpine catchments of the Cordillera Blanca, Peru. However, tropical alpine glaciers in this region are melting rapidly, and as they progress past peak water, their ability to sustain dry season streamflow will continually decrease. This study sought to improve our understanding of the groundwater hydrology within the pampa aquifers of such headwater catchments since groundwater is the other main contributor to streamflow during the dry season. We created a groundwater flow model of a portion of a proglacial pampa aquifer system in the Quilcayhuanca Valley of the Cordillera Blanca to investigate the relative groundwater contribution to streamflow and the amount of time required for recharge to travel through the aquifer system to the stream. The groundwater system in these valleys consists of a confined aquifer made of glaciofluvial and glaciolacustrine deposits, intermixed with rockfall deposits. Hydraulic head data from six piezometers and estimates of groundwater flux to the stream from previous heat and dye tracing and water chemistry end-member mixing analysis studies were used to calibrate the groundwater flow model. Model results indicate the precipitation that falls on the cliff faces and talus slopes enters the confined valley aquifer from the sides, and is the dominant source of aquifer recharge; the precipitation that falls on these slopes takes about two months to move through talus deposits on the valley periphery and reach the aquifer. The portion of streamflow the model estimates to originate as groundwater during July (~37%) within this catchment is comparable to previous July estimates (38-53%). Model results also show that the

travel time of groundwater through this aquifer system is relatively short (<1.5 years), increasing the region's vulnerability to future dry periods.

1 Introduction

The water supply of at least one-sixth of the world's population is dependent on melt water from glaciers and seasonal snowpack (Barnett et al., 2005). Tropical glaciers, which occur in the equatorial region between 23.4°N and 23.4°S at high elevations, are an important source of streamflow in dry or seasonally dry regions. Over 99% of the world's areal extent of tropical glaciers are located in South America and ~71% of tropical glaciated areas are located in Peru (Kaser, 1999). The highest density of tropical alpine glaciers occurs in the Peruvian Andes in the Cordillera Blanca (Kaser et al., 2003). Unfortunately, the snow and ice melt processes in these watersheds are highly vulnerable to climate change (Barnett et al., 2005). Although Peru contains the highest density of tropical glaciers worldwide, water resources in the region are limited, especially during the dry season when only ~20% of annual precipitation occurs (Burns et al., 2011). While glacial melt water currently helps sustain dry season streamflow in Peru (Mark & Seltzer, 2003), the loss of alpine glaciers due to climate change will eliminate the buffering capability provided by glacial melt. This is particularly concerning because the glaciated valleys of the Cordillera Blanca drain into the Rio Santa, which is an important regional water source for agriculture, mining, hydroelectric and other industries, in addition to providing water to the local communities (Bury et al., 2013). Many glaciers have already passed 'peak water', and as a result, discharge from these watersheds will decrease to a lower equilibrium level that ultimately lacks glacial inputs (Bury et al., 2013). Some studies estimate up to a 60-70% decrease in dry season discharge in the most vulnerable watersheds, along with increasing seasonal discharge

variability, further stressing Peru's limited water resources (Mark & Seltzer, 2003; Bury et al., 2011; Baraer et al., 2012).

However, recent research indicates that groundwater stored within alpine systems can have an important and previously overlooked contribution to streamflow (Clow et al., 2003; Hood et al., 2006). In particular, sediment deposits including talus units and moraines within alpine watersheds can have an important influence on groundwater storage potential, along with how and when groundwater is released to alpine streams (Roy & Hayashi, 2009; McClymont et al., 2011; Muir et al., 2011; McClymont et al., 2011). Glaciated catchments within the Cordillera Blanca contain many such deposits that likely store wet season precipitation and release it to the streams during the dry season. The high elevation wetland meadow systems in the catchments, known as pampas, are made of clay and organic rich surficial deposits that overlie talus/landslide and moraine deposits that are infilled with glaciolacustrine and glaciofluvial sediments (Chavez, 2013; Glas 2018). Prior research indicates these complex meadow systems store a portion of the wet season precipitation and release the water to streams via a combination of springs and subsurface flow during the dry season (Baraer et al., 2015; Gordon et al., 2015). These studies estimate that anywhere from 24-80% of streamflow originates as groundwater during the dry season, depending on the extent of glacial coverage within the catchment (Baraer et al., 2015; Somers et al., 2016). However, the resiliency of these groundwater resources is uncertain since little is known about the residence time distribution of groundwater within these Andean alpine meadow systems. One study estimates bimodal groundwater residence times of ~3 months and ~1.5-3 years in one Peruvian proglacial valley based on the correlation between groundwater contributions derived from a mixing model and antecedent precipitation (Baraer et al., 2009). Additional research is therefore needed to provide a more thorough understanding of the travel

time distributions of groundwater within proglacial valleys of the Cordillera Blanca to determine how resilient this alpine groundwater resource may be to regional water stresses. In this study we model groundwater flow through an alpine pampa aquifer system in the Cordillera Blanca that contains a headwater stream of the Rio Santa. This groundwater flow model is then used to (1) estimate the amount of groundwater entering the stream both diffusely and through springs, (2) estimate groundwater residence time, and (3) ultimately determine if these alpine meadow aquifer systems store enough wet season precipitation to sustain dry season stream flows.

2 Study Site & Conceptual Model

The Cordillera Blanca, Peru, has ~631 km^2 of glacial coverage, making it the most glacierized mountain range in the tropics (Suarez et al., 2008). The region experiences distinct wet and dry seasons due to the seasonal oscillation of the Intertropical Convergence Zone, with over 80% of precipitation occurring between October and May (Kaser et al., 1990; Kaser et al., 2003; Bury et al., 2011). Runoff from the Cordillera Blanca drains to the Rio Santa, which flows west to the Pacific Ocean across the arid western region of Peru. While precipitation is highly seasonal, annual air temperatures vary minimally, with the daily air temperature fluctuating more than the average annual air temperature (Kaser et al., 1990). This causes glacial ablation to occur throughout the year, rather than seasonally, with greater glacial ablation occurring during the wet season than the dry season (Kaser et al., 1990; Mark & Seltzer, 2003; Suarez et al., 2008), and providing melt water to streams within these glaciated catchments.

The study site is located within the Quilcayhuanca valley in the Cordillera Blanca and is typical of proglacial valleys in the northern portion of this mountain range (Figure 1). The Quilcayhuanca basin is approximately 88 km^2, ~20.5% of which is glaciated (Baraer et al., 2015). The majority of glacial area is found in the highest elevation headwater regions of the

valley (Glas et al., 2018). The valley bottom consists of pampa wetlands (~4 km^2) that are occasionally crosscut by landslides and moraines and provide groundwater discharge to springs and the stream (Glas, 2018). Pampas are high elevation wetland meadow systems covered in grass and consist of glaciolacustrine sediments and organic material that formed from paludified glacial lakes (Mark & McKenzie, 2007; Vincent et al., 2019). The sides of the valley are steep granodiorite cliffs, the lower portions of which are covered in talus and rock-fall deposits. The bedrock in the valley is primarily granodiorite, though pyrite-rich metasedimentary rock of the Chicama formation is exposed in the uppermost portion of the valley. The sediments within the valley consist of glaciolacustrine and glaciofluvial sediments, along with interbedded rockfall deposits at depth and a layer of peat and organics at the surface (Figure 2). The surficial deposits consist of silty clays intermixed with sand, which act as a confining layer; these deposits were formed at the time a glacial lake filled the valley. A confined aquifer is found below the surficial confining layer and consists of a mix of boulders, sand, and gravel (Chavez, 2013).

The model domain contains a reach of the Quilcay stream approximately 0.89 km long. The domain begins ~0.23 km below the base of a steep debris structure (moraine or talus deposit) and encompasses a portion of the low gradient pampa. Upstream of the model domain, the Quilcay stream flows through additional pampas and debris structures and is sourced from glacial lakes and direct glacial melt. The majority of the stream reach in the model domain is single threaded, with one small braided reach near a gravel bar. Stream flow can exceed 9.0 m^3/s in the wet season and fall below 0.6 m^3/s during the dry season; streamflow is highest in February with an average of ~3.4 m^3/s and lowest in July with an average of ~1.0 m^3/s (Figure 3). The pampa within the model domain contains many small springs and tributaries, some ephemeral and some perennial. Talus and landslide deposits border the edges of the pampa, at

the base of the cliffs. The surface of the wetland meadow (pampa) is hummocky and covered in short grazed grasses and organic material.

3 Methods

A range of methods were used to collect the field data required to construct the groundwater flow model within the Quilcayhuanca catchment. Additional data were also compiled from previous research conducted in the catchment. A transient groundwater flow model was then created to simulate the annual variability of groundwater levels in the model domain and estimate groundwater residence time. Various model configurations were explored to determine which best simulated the observed average seasonal hydraulic head fluctuations within the aquifer and the previously estimated groundwater fluxes to the Quilcay stream.

3.1 Field Methods

Seven boreholes were drilled across the pampa to install piezometers screened in the aquifer (Figure 1; Table 1). The top 0.61 m to 6.20 m of each borehole consisted of silty clay mixed with medium grained sand of glaciolacustrine origin, which acts as a confining layer of the underlying aquifer (Chavez, 2013). Below the confining layer is a heterogeneous layer of gravel and boulders mixed with sand and silt, forming the aquifer (Chavez, 2013); rockfall deposits from the cliffs also crosscut the valley deposits. Seismic data and 2D resistivity surveys indicate the depth to a more conductive (aquifer) material is likely 2 m to 8 m below the ground surface (Glas, 2018). Vertical electrical sounding surveys show a resistivity increase that indicates bedrock or clay rich till occurs 18 m to 35 m below the ground surface, marking the base of the aquifer. Horizontal-to-vertical spectral ratio data indicate the depth to competent bedrock is 20 m to 85 m below the ground surface (Glas, 2018). Though geophysics data cannot

determine exact unit boundaries, it constrains a range of aquifer and confining layer thicknesses to test in the groundwater flow model.

Groundwater levels (corrected for atmospheric pressure) were measured every hour using Schlumberger Mini-Diver pressure loggers (Chavez, 2013) installed in the six piezometers screened within the aquifer (Figure 4). One additional borehole was artesian at the time of drilling and therefore had to be plugged (Chavez, 2013). The piezometers are oriented such that one transect is perpendicular to the stream and the other is oriented down valley, providing information on both the cross-valley and down-valley slope of the piezometric surface (Figure 1). Slug tests were performed in three of the piezometers to estimate the hydraulic conductivity (Table 1) of the aquifer (Chavez, 2013). Hydraulic conductivity measurements of the aquifer ranged from 6.5×10^{-6} m/s to 7.3×10^{-5} m/s. While no hydraulic conductivity measurements were made of the surficial unit, we know it has a lower hydraulic conductivity since it acts as a confining layer. Other hydrologic parameters (S_y, n_e, S_s; Table 2) were estimated from known values of similar materials (Smith & Wheatcraft, 1993).

Stream stage was measured every 15 minutes (July 2009 to February 2014) at a gauging station in the approximate center of the study reach and converted to discharge using a rating curve (Figure 3). A meteorological station located near the gauging station recorded precipitation data every hour from July 2013 to July 2016 (Figure 5). Over the three-year period, the average precipitation rate was 776 mm/year.

3.2 Groundwater Flow Modeling

Groundwater flow modeling was performed using Visual MODFLOW Flex 4.1 (Waterloo Hydrogeologic Inc., 2005) which is a graphical user interface that runs the USGS MODFLOW-2005 finite-difference computer code that iteratively solves the groundwater flow

equation (Harbaugh, 2005). We developed a transient model of the study site to simulate the seasonality of groundwater levels caused by the distinct wet and dry seasons that occur in the region. Our three-dimensional model domain was divided into 68 rows and 49 columns in order to create 10 m by 10 m grid cells. The model domain encompasses a pampa region with an area of ~0.09 km^2. The Quilcay stream forms a natural hydrologic boundary, and only the area on the southeastern side of the stream was modeled. The extent of the pampa on the northwestern side of the stream is a small percentage of the overall area and lacks any observation wells to inform model calibration (Figure 1). The model consists of two hydrostratigraphic units; the top confining layer consists of one cell layer, while the bottom aquifer unit is divided into three cell layers. A 10 cm resolution DEM of the pampa (Wigmore & Mark, 2018) was resampled to 5 m resolution and used as the surface topography (top of confining layer) in the model. The elevation of the top of the aquifer unit was set to either 3 m or 4 m below the top of the model. The base of the model was set as either 15 m or 30 m below the surface of the model. Four different model domain configurations were simulated to encompass our uncertainty in the thicknesses of the hydrostratigraphic units (Figure 6). Confining layer thickness is based on borehole (Chavez, 2013) and geophysics data (Glas, 2018), while aquifer thickness is based solely on the geophysics data.

Multiple types of boundary conditions were used in the model. The stream is modeled as a specified head boundary, and its location coincides with about half of the model perimeter. To determine the elevation of the specified head boundary, the average stage recorded at the gauging station was first calculated for each month, using ~4.5 years of stage data. The streambed elevation along the length of the stream was measured using the 10 cm DEM. Then the calculated average stage value for each month was added to the streambed elevation to

determine the elevation of the stream specified head boundary during each month (Figure 3). Drains were assigned to the uppermost layer to simulate ephemeral groundwater springs that vary seasonally. Drain conductance values were derived by assuming the hydraulic conductivity of the drain bed is equal to the hydraulic conductivity of the aquifer unit. The southwest side of the model domain is a specified flux boundary to simulate groundwater flux from the talus slopes into the pampa aquifer. A small amount of recharge was also applied to the surface of the model domain based on the seasonal variation in precipitation (Figure 5).

Transient simulations were run by repeating average monthly conditions for a 16-year period. By repeating the same average monthly conditions for multiple years, the model achieved a dynamic steady state representation of long-term seasonality. The initial heads for the transient simulations were the heads from the last corresponding stress period of a prior ~16-year transient model run. Twelve stress periods were applied for over the year, each representing a month of time. The inputs for the boundaries all varied through time, using monthly, long term averages (Figure 3, Figure 5). For the specified head boundary (the Quilcay stream), the stage value for each monthly time step was calculated as the average stream stage during that month over the ~4.5-year observation period at the gauging station (Figure 3a-b). The direct surface recharge and groundwater influx from the side talus slope deposits (specified flux boundary) were both scaled according to average monthly precipitation values over the 3-year observation period at the meteorological station. The maximum possible recharge values were determined based on the amount of average monthly precipitation. The amount of direct surface recharge applied to the pampa was half of the monthly precipitation rate (50%). For the groundwater flux from the side talus slopes (specified flux boundary), a maximum volume of water was calculated based on the average monthly precipitation rate and the combined talus (~0.3 km^2) and cliff slope (~1.75 km^2)

surface areas whose runoff is assumed to infiltrate the talus slope deposits and recharge the aquifer from the sides; a percentage of that maximum flux value was then used in the model based on calibration results. This recharge from the side talus slopes was lagged by two months to account for the travel time through the talus deposits. For some model runs, we assumed a greater percentage of the precipitation volume infiltrated the talus slopes during the driest months than during the wettest months, because we assumed the storage capacity of the talus was exceeded during the wettest months. For other model runs an additional lag was incorporated to account for variable distances that groundwater may have to travel within/along the cliff face and talus slope; 50% of the precipitation had a two-month lag, 30% of precipitation had a three-month lag, and 20% of precipitation had a four-month lag.

3.3 Model Calibration

The groundwater flow model was calibrated using both hydraulic head observations and estimates of groundwater discharge to the stream and springs within the modeled area. Monthly averages of observed hydraulic head data from the six piezometers within the model domain were compared to the hydraulic head values calculated by the model; data within 5 cm of either the base of the piezometers or the casing elevation were excluded from the calibration to ignore periods when the piezometers dried out or when water levels exceeded the height of the land surface. Both qualitative and quantitative comparisons were made to assess how well the model reproduced average monthly observed hydraulic heads. Quantitatively, the head values were compared using error metrics such as the root mean squared error (RMSE). The RMSE for each model was calculated as the average of the RMSE of all piezometers. Additionally, for each month, the slope of the average measured versus modeled heads was calculated. A slope of 1 represents a match of the observed and modeled heads. The average R^2 of the modeled versus

observed heads for each month was also calculated. Qualitatively, modeled head values should exhibit the same seasonality as observed data, with some piezometers (P3) drying up during the dry season and all piezometers overtopping the casing/over pressurizing during the wet season.

The models were also calibrated by comparing the groundwater discharge to the reach estimated in other Quilcayhuanca studies (58.6 L s^{-1} km^{-1} in Somers et al., 2016; 42.1 L s^{-1} km^{-1} in Baker et al., 2018) to the amount of groundwater exiting the model domain through the stream (specified head boundary) and springs (drain boundary), thereby ensuring a reasonable amount of groundwater was exiting the model via these pathways. These studies estimate that springs and groundwater contribute ~5-8% of streamflow within the model reach. Estimates of the spring and groundwater influx within the reach during the dry season were also made using simple mixing equations of the major ion water chemistry data (Figure 7) and compared to the calculated fluxes out of the model domain. The chemistry data indicate a spring and groundwater influx to the reach of ~3-21% of streamflow during the dry season (Table 3). However, flux rates out of the model via springs and groundwater will be less than these observed values because the model only encompasses the pampa on the southeastern side of the river. Additionally, the largest spring in the pampa is found on the northwestern side of the stream, and is not included in the model domain. Therefore, the total groundwater flux calculated by the model was multiplied by a factor of two to account for its simulation of the pampa on only one side of the stream. For each model run, the total flux of water out of the model into the springs (drain boundaries) and stream (specified head boundary) was determined using ZoneBudget in MODFLOW; the values for July and August are reported (Supplementary Table S1) since that is when water chemistry, heat budget and dye tracing analyses were conducted to estimate the dry season flux of groundwater into the stream.

3.4 Sensitivity Analysis

We assessed the model sensitivity to various model input parameters over our inferred range of uncertainty. Simulations were run varying the confining layer thickness (3 m, 4 m) and the aquifer thickness (15 m, 30 m, 50 m). The confining layer hydraulic conductivity was varied from $1x10^{-7}$ m/s to $1x10^{-4}$ m/s in the model simulations, while the aquifer hydraulic conductivity was varied from $1x10^{-5}$ m/s to $1x10^{-3}$ m/s. We also varied the anisotropy of the hydraulic conductivity. Sedimentary deposits can have hydraulic conductivity anisotropy ratios of 2:1 to 10:1, and at a larger scale anisotropy can be over 100:1 (Smith & Wheatcraft, 1993). Some model runs incorporated no hydraulic conductivity anisotropy, some runs had a ratio of 5:1, and some runs had a ratio of 10:1. The recharge from the side talus slopes was varied from 30% to 70% of the volume of precipitation that fell on the talus and cliffs, while the direct surface recharge was varied from 10% to 50%; some simulations were also run where only the volume of precipitation that fell on the talus portion of the slopes were varied (volume on cliffs was excluded). The drain conductance and the storage properties were also varied (S_y, n, S_s). The parameter combinations for a subset of model simulations are in Supplementary Table S1 and the results of the sensitivity analysis are in Table 4.

3.5 Groundwater Residence Time

During each model run, particles were forward modeled to estimate the travel time required for groundwater to move through and exit the meadow aquifer system. A group of ~270 particles was released into the model at all depths along the cells that received the side flux from the talus slope recharge at the beginning of the 16-year transient model run. Both the travel time and the distance traveled by all the particles were calculated in MODFLOW using MODPATH.

3.6 Sensitivity to Changing Environmental Conditions

The model with the best match to the observed conditions based on the various criteria used to assess model performance was then used to assess how changes in the amount of recharge to the aquifer would affect groundwater flow within the pampa. Model simulations were run where the monthly precipitation rates were reduced by 5% or 10% for a one-year period beginning at the start of the wet season, to assess how the hydraulic heads and groundwater fluxes would be affected by a period of drought. These simulations can also provide insight into how hydraulic heads and groundwater fluxes would change if recharge decreased due to higher air temperatures (more evaporation) and/or lower precipitation rates.

4 Results

4.1 Model Simulation of Groundwater Heads

Numerous model configurations were run to improve the model fit to the observed hydraulic heads observed in the piezometers throughout the valley. If the hydraulic conductivity of the aquifer is larger than $1x10^{-3}$ m/s, there is not enough precipitation available to recharge the aquifer and maintain the magnitude or the slopes of the hydraulic heads. If the hydraulic conductivity of the confining layer is two orders of magnitude lower than the aquifer hydraulic conductivity, the aquifer becomes extremely over pressurized. Additionally, if the hydraulic conductivity of the confining layer is less than $1x10^{-7}$ m/s, the heads in the aquifer are too high and can only be brought down to a realistic level if the recharge is decreased; however, this results in groundwater fluxes to the stream that are less than estimated by other studies.

Modeled groundwater heads within the confined aquifer were typically overestimated during the wet season and underestimated during the dry season (Figure 8). Modeled heads during the dry season were closest to observed heads at piezometers P5 and P6, which were located closest to the stream (Figure 8d-e). The modeled cells in the confining layer often dry out

during the dry season, particularly in piezometers P3 and P7, which is consistent with field observations. The slope of the piezometric surface is more accurately modeled during the months at the end of the wet season and beginning of the dry season (April to July) than during the months at the end of the dry season and beginning of the wet season (August to November; Supplementary Figures S1-S2). Additionally, the modeled heads in layer 2 are often slightly higher than the modeled heads in layer 1 (Supplementary Figures S1); these artesian conditions within the aquifer unit are supported by observed artesian conditions in the field.

The lowest average monthly RMSE of the hydraulic heads occur when only precipitation that falls on the talus slopes contributes to aquifer recharge through the side talus slopes, and precipitation that falls on the upper portions of the cliffs is excluded (Run A: RMSE = 1.13 m). However, the total groundwater contribution within the simulated model reach for these configurations is too low (<1% in July). Therefore, in order to increase the total groundwater flux to the stream, the precipitation that falls on both the cliff and talus slopes must be considered, and the hydraulic conductivity of the aquifer and confining layers must be increased. Runs C and D have the lowest average hydraulic head RMSE of all the model configurations run with aquifer recharge contributed by the total talus slope/cliff face area (Supplementary Table S1). The six piezometers had an average hydraulic head RMSE of 1.21 ± 0.73 m in run C and an average RMSE of 1.28 ± 0.90 m in run D. Both of these model configurations had 27 m thick aquifer units with a horizontal hydraulic conductivity (K_h) of $5x10^{-5}$ m/s, with a K_h/K_v of 5:1. However, the hydraulic conductivity of the confining layer was an order of magnitude larger in run D. For run C the side talus slope/cliff face recharge rate was set to 30% of the monthly precipitation rate during the wet season, but 70% of the monthly precipitation rate during the driest months, with a two-month lag. For run D, 50% of the monthly precipitation that fell on the

talus slope/cliff face was set to recharge the aquifer, but 50% of the infiltrating precipitation had a one-month lag, 30% had a two-month lag, and 20% had three-month lag.

Runs R and T had the strongest monthly correlation between the measured and modeled heads of all the model configurations run with aquifer recharge contributed by the total talus slope/cliff face area (Supplementary Table S1). For any month, a slope of one between the modeled and measured heads indicates a perfect correlation. Run R had an average monthly correlation of 1.01 ± 0.09 and run T had a correlation of 0.99 ± 0.17. These runs had higher monthly average RMSE values of 1.43 m ± 0.9 m and 1.49 m ± 0.91 m. The slope of the piezometric surface was most accurate in June and July in run both runs R and T (Supplementary Figures S1-S2). In addition to the RMSE and monthly head correlations, we also assessed the dry season groundwater fluxes to determine which model configuration is a more accurate representation of the pampa aquifer system.

4.2 Groundwater Contribution to Streamflow

The flux of groundwater into the springs and stream were estimated using the flux into the springs (drain cells) and stream (specified head cells). The flux to these cells during July and August were used to help with model calibration since the heat tracing, dye tracing, and hydrochemical mixing modeling used to estimate groundwater flux to the stream were conducted during this time period. The model with the lowest RMSE had fluxes to the drain and spring cells totaling 1151 m^3/day/km (Run C) in July. Since the model only encompasses the aquifer system on one side of the stream, and the biggest spring occurs on the other side of the stream, we estimate that these flux values are only half of the total groundwater flux to the stream within the model domain. Therefore, run C indicates a gain in streamflow of ~2.6% due to total groundwater influx. Meanwhile, the simulation whose modeled and measured heads during any

month have the strongest correlation estimates groundwater contributes 5.0% (Run T) of streamflow along the model reach in July (Supplementary Figure S3a). If only precipitation that falls on the side talus slopes is included, rather than also including precipitation that falls on the cliff face, then the groundwater flux is lower than estimated with previous methods (e.g. Run A: 1.0% influx). Previous dye and heat tracing studies estimated the groundwater gain along this reach was ~5-8% of stream flow (Somers et al., 2016; Baker et al., 2018), while previous hydrochemical mixing analyses and simple water chemistry mixing estimated that ~6-11% (Baraer et al., 2015) and ~3-21% (Table 3) of water in the model reach was from groundwater entering within the reach. Therefore, the July fluxes estimated by run T agree most closely with the groundwater gains estimated by other methods conducted during this time, while also meeting the requirement of a strong monthly head correlation, even though the RMSE was larger than some of the other simulations. Due to this, run T was chosen as the model configuration that best represents the pampa aquifer system.

4.3 Model Sensitivity

The calculated hydraulic heads and average total groundwater fluxes from pairs of simulations were compared to assess the sensitivity of the models to the input parameters. If the model thickness is increased from 15 m to 30 m (aquifer thickness increased from 12 m to 27 m), the heads decrease by an average of 0.25 m and the total flux to the stream increases by an average of 297 m^3/day (9.7% of mean flux). Increasing the confining layer thickness from 3 m to 4 m only increased heads by 0.18 m and decreased the groundwater flux by 0.8%. If the confining layer hydraulic conductivity is increased by one order of magnitude ($1x10^{-5}$ m/s to $1x10^{-4}$ m/s), the heads decrease by an average of 0.86 m and the total flux to the stream increases by an average of 114 m^3/day (3.4% of mean flux). Meanwhile, increasing the aquifer hydraulic

conductivity by only half an order of magnitude produced a comparable decline in head of 0.75 m, and a groundwater flux change of 3.4%. Increasing the anisotropy ratio of the layers (the difference between the horizontal and vertical hydraulic conductivity values) from 5:1 to 10:1 results in a 2.5 m increase in head and a decrease in groundwater flux of 2.5%. The direct surface recharge to the pampa was varied but did not have much of an effect on the heads observed in the aquifer; reducing surface recharge from 50% of the precipitation to 10% only reduced heads by 6 cm and reduced the average groundwater flux by 5%. This may be because much of this recharge never reaches the aquifer but rather exits the model as spring flow, or may be because this volume of recharge is small relative to the recharge from the side talus slopes. If the side talus slope recharge increases from 50% to 70% of the precipitation falling on the cliff/talus slopes, the heads decrease by an average of 0.32 m and the total flux to the stream decreases by an average of 841 m^3/day (37.9% of mean flux). Decreasing the drain conductance by two orders of magnitude results in an increase in head of 1.5 m and a negligible change in groundwater flux. Decreasing specific storage (S_s) values by an order of magnitude and decreasing porosity and specific yield values by 0.1 all had produced minimal changes to the average hydraulic heads and groundwater fluxes calculated by the model (Table 4). Overall the uncertainty in the anisotropy ratio of the sediments produces the largest change in the calculated hydraulic head, while the uncertainty in the valley side recharge produces the largest change in the calculated total groundwater flux (Table 4).

4.4 Groundwater Travel Time Distribution

Approximately 270 particles were forward modeled using MODPATH and released into the model through the side talus slope recharge flux. The travel times of particles through the aquifer system were relatively short for all model configurations that used recharge from the

combined talus slope/cliff area; the longest maximum residence time of these simulations was ~4 years (Table 5). Run T, which was chosen as the most representative model configuration, estimated a maximum travel time of 1.5 years, and a median travel time of 85 days (Table 5). The majority of particle travel times (84%) are less than 6 months (Figure 9).

4.5 Impact of Decreased Recharge

Observations of air temperatures in the tropical Andes show that mean annual air temperatures are increasing at rates of 0.1-0.3°C per decade (Schauwecker et al., 2014; Vuille et al., 2008; Vuille, Kaser & Juen, 2008), with a rate of 0.31°C per decade from 1969 to 1998 and 0.13°C per decade from 1983-2012 in the Cordillera Blanca (Schauwecker et al., 2014). Meanwhile, no clear precipitation trends are present in the Cordillera Blanca (Glas et al., 2018). Precipitation data recorded from July 2013 to July 2016 show up to 30% variability in the annual precipitation rate (Figure 5). Model simulations were run with lesser recharge than the best model configuration to explore the effect of a drier and/or warmer climate in the region. If the recharge is decreased by 5% for only a one-year period, beginning at the start of the wet season, the hydraulic heads in the piezometers decrease by an average of 7.9 cm and the total groundwater flux decreases by an average of 4.4%. If the recharge is decreased by 10% for only a one-year period, beginning at the start of the wet season, the hydraulic heads in the piezometers decrease by an average of 16 cm and the total groundwater flux decreases by an average of 9.2%. Months at the end of the dry season (September to November) experienced the smallest change in hydraulic heads (Figure 11a) and groundwater fluxes (Figure 11b). The effect of a drought does not become evident until a few months after it has begun, due to the travel time required for precipitation falling on the talus and cliffs to reach the aquifer.

5 Discussion

5.1 Which fluxes drive observed groundwater heads in pampa aquifers?

Modeling the seasonal groundwater heads in the Quilcayhuanca confined aquifer has refined our conceptual understanding of groundwater recharge and discharge mechanisms in these proglacial valley sediments. We found that even if all precipitation falling on talus slopes on the valley periphery infiltrates and ultimately recharges the valley bottom aquifer, the amount of recharge is not sufficient to maintain both the measured hydraulic heads and the estimated groundwater fluxes to the stream. Rather, all the precipitation that falls on the entire cliff and talus slope system draining to the valley bottom must be considered. Therefore, the total amount of precipitation that falls on the combined talus slope/cliff face surface is an important source of recharge to these confined meadow aquifer systems. This means that although the pampas comprise only a small percentage of the catchment area, they receive water draining from a march larger portion of the catchment and have potential to store these large volumes of water for later delivery to the stream (e.g. during the dry season).

Furthermore, we found that the flux of water from the cliff and talus slope must be lagged ~2 months behind when the precipitation occurs to obtain the correct seasonality in groundwater heads. This lag time helps account for the distance water has to travel through the talus slope deposits before entering the pampa aquifer. We also found that the modeled monthly head correlations and estimated groundwater fluxes improve if a greater percentage of the precipitation infiltrates the subsurface during the dry season months. This may be due to the storage capacity of the talus slopes being exceeded during the wet season months; the talus slopes and pampa aquifer fill quickly at the onset of the wet season and then the remaining precipitation runs off rather than infiltrating the talus slopes.

5.2 What is the residence time of groundwater in a proglacial catchment?

While none of the models completely captured the annual variation in the magnitude of the groundwater heads, estimates of the groundwater residence time were consistent regardless of the model configuration, with ~4 years as the longest maximum travel time of groundwater particles. The model that was chosen as most representative of the meadow aquifer system predicted a median travel time through the aquifer of ~2.8 months, with a maximum travel time of 18 months and a subset of particles with travel times ranging from ~12 to 18 months (Figure 9). This finding agrees with groundwater residence times estimated in the Querococha watershed, which is a few valleys south of Quilcayhuanca. Using antecedent precipitation analysis, a prior study estimated a maximum residence time of 4 years and shorter groundwater flow paths that resulted in ~3-month and 18 to 36-month residence times (Baraer et al., 2009). Although the maximum residence time predicted by this model is only 1.5 years, the longer flow paths of ~4 years predicted by antecedent precipitation analysis could occur deeper in the aquifer where hydraulic conductivity values are likely smaller, though the model simulations assign a homogeneous hydraulic conductivity to the whole aquifer unit. The similarity between the residence times predicted by these two methods gives us confidence that this groundwater flow model is representative of the pampa aquifer system. In addition, the similarity between the travel times predicted for these two proglacial valley aquifer systems indicates the pampa aquifer systems in these adjacent valleys and throughout the individual valleys likely behave similarly.

5.3 How much groundwater enters proglacial streams?

The amount of groundwater that enters the stream from groundwater discharge in the catchment above the gaging station can be estimated. To estimate the amount of groundwater that has entered the Quilcay stream throughout all pampa regions upstream of the gaging station

(Figure 1), we multiplied the groundwater flux rate to the model reach for each month by the total length of stream that flows through pampa regions (7.34 km), and then doubled this value to account for the model domain only simulating one side of the pampa aquifer system. This rate was then converted to a percentage of the total streamflow at the gaging station for each month (Figure 10b). During July, the estimated total relative groundwater contribution is ~37%, which is similar to the 38-52% range predicted previously by end-member mixing analyses performed on water chemistry data sampled during July 2008 (Baraer et al., 2015). Additional groundwater also enters the stream where it flows through moraine and landslide deposits (Figure 1a), however lower flux rates to the stream are estimated in these regions of the catchment. For example, previous stream temperature energy balance modeling estimated a groundwater flux into the stream at a rate of 6.5 L/s/100 m within meadow aquifer regions, and an influx rate of 4.0 L/s/100 m in moraine/landslide reaches of the stream (Somers et al., 2016). Therefore, these reaches contribute less groundwater to the stream, both due to the lower flux rate and their lesser areal extent. However, inclusion of this source of groundwater to the groundwater contribution estimated in this study could raise the total groundwater contribution to stream flow to the upper end of the range predicted by end-member mixing analyses.

The groundwater flux to the stream represents the largest percentage of streamflow at the onset of the dry season (May-July) (Figure 10b). During this period, the total groundwater contribution (direct + springs) to stream flow ranges from ~37-53% (Figure 10), while the direct groundwater contribution to stream flow ranges from 16-17% (Figure 10b). Direct groundwater flux contributes the largest percentage during June and July because streamflow is declining as less precipitation occurs, but the aquifer is still receiving lagged recharge from the talus slopes/cliff face and contains an excess of water that drains out into the stream. The direct

groundwater flux to the stream is lowest in October and November at ~4-5% of streamflow, at the start of the wet season. During this time, the streamflow is increasing due to increased precipitation, while the direct groundwater flux to the stream is small; the aquifer is depleted at the end of the dry season and wet season precipitation has not yet traveled through the talus slopes to recharge the aquifer. At the beginning of September (end of dry season), before precipitation begins to increase at the onset of the wet season, ~13% of streamflow originates as groundwater, with ~9% of streamflow from direct groundwater flux (Figure 10b). We also see that the groundwater springs can contribute substantially to stream flow, depending on the time of year. For example, from February through July, stream flow consists of 17-38% spring water (Figure 10b). In general, the relative total groundwater contribution to the stream is largest during May and June, while the relative direct groundwater contribution is largest during June and July. The relative groundwater contribution to the stream declines over the course of the dry season, reaching a minimum in October. While different model configurations vary on the exact amount of total groundwater entering the stream, the annual pattern of groundwater discharge to the stream is similar across all simulations (Supplementary Figure S3b).

5.4 Model Limitations

Many of the limitations of this model come from uncertainty in certain model parameters. For example, half an order of magnitude difference in the anisotropy of the confining layer and aquifer can influence hydraulic head values by an average of ~2.5 m. Sensitivity of the model to the amount of recharge fluxing from the side talus slopes into the aquifer also leads to model uncertainty as a 20% increase in the volume of precipitation that infiltrates the talus slopes and recharges the aquifer produces an average hydraulic head difference of 0.3 m and an average change in groundwater flux of 38%. Although we do not precisely know some of the aquifer

storage parameters (S_s, n_e, etc.), the difference in head and groundwater fluxes produced by our uncertainty in their true value is relatively small (Table 4). Similarly, we have found that even though the aquifer thickness is uncertain, doubling the thickness only decreases hydraulic heads by 0.25 m and increases the average amount of groundwater flux by 10% (Table 4). Though approximately doubling the aquifer thickness can almost triple the particle travel time, the estimated groundwater travel times are short enough that this does not have a large effect on the overall conclusions (Run U vs. Run W; Supplementary Table S1). While additional data on aquifer thickness, confining unit hydraulic conductivity, unit anisotropy, and aquifer recharge rates could all help to better constrain the model, the site data and our conceptual model of the groundwater system produce a model that seems to accurately represent the system given our current knowledge. Considering the uncertainty in certain model inputs, the coarseness of the model grid, the use of average monthly data from only partially overlapping observation periods, and the simplification of the heterogeneous subsurface units into two homogenous units, the modeled hydraulic heads and groundwater fluxes match the observed and estimated values quite well. Therefore, even though this model is a very simplified version of the proglacial pampa aquifer system, we have confidence in the model results since they align well with the observed head data and the estimates from previous studies that made use of other methods to estimate groundwater contribution to streamflow and groundwater residence time.

5.5 Aquifer Vulnerability to Dry Periods

When the amount of recharge entering the model through the side talus slope flux decreases for a year due to a 5-10% decrease in the annual precipitation rate, the elevation of the piezometric surface decreases and the groundwater flux out of the aquifer and into the stream decreases proportionally (Figure 11Xb). Therefore, while groundwater does contribute to

streamflow, and substantially to streamflow during certain months of the year, it can only temporarily help to sustain streamflow during dry periods as the maximum travel time of groundwater particles through the system is only about 1.5 years. These short travel times make the pampa aquifer systems vulnerable to droughts. An extended drought during the wet season can decrease the groundwater contribution to streamflow in the following dry season (Figure 11b). Additionally, streamflow at the end of the dry season is the most vulnerable to a reduction in glacial meltwater since groundwater and precipitation constitute a small relative contribution to streamflow during this time (Figure 10b).

6 Conclusions

By combining a variety of field data and observations, we were able to create a groundwater flow model that accurately simulates hydraulic heads and groundwater fluxes within a proglacial alpine catchment of the Peruvian Andes. About 7-53% of Quilcayhuanca streamflow during the dry season is derived from groundwater, with ~5-17% derived from direct groundwater inflow depending on the month; this groundwater contribution declines throughout the dry season as the groundwater reserves become depleted. The model also reveals that groundwater contribution to streamflow from direct groundwater flux and via springs are both important. Groundwater within this proglacial pampa aquifer system has relatively short residence times, with >80% of particles that enter as recharge from the side talus slope/cliff face exiting the model in 6 months. The remaining particles travel through the aquifer in around 1-1.5 years. We also determined that precipitation that falls on the talus slopes and cliff faces is critical to recharging the aquifers in these systems, and that this precipitation takes ~2 months to travel from where it lands on the valley sides down into the aquifer through the talus slope deposits. These short travel times imply that such alpine pampa aquifer systems in the Andes are sensitive

to seasonal to yearly variations in precipitation, and that stream flow through such systems is vulnerable to drought. Lastly, months at the end of the dry season are most vulnerable to glacial loss, since the groundwater contribution to streamflow is smallest during this time.

References

Baker, E. A., Lautz, L. K., Kelleher, C. A., & McKenzie, J. M. (2018). The Importance of Incorporating Diurnally Fluctuating Stream Discharge in Stream Temperature Energy Balance Models. *Hydrological Processes*, 32, 2901–2914. https://doi.org/10.1002/hyp.13226

Baraer, M., McKenzie, J. M., Mark, B. G., Bury, J., & Knox, S. (2009). Characterizing contributions of glacier melt and groundwater during the dry season in a poorly gauged catchment of the Cordillera Blanca (Peru). *Advances in Geosciences*. http://doi.org/10.5194/adgeo-22-41-2009

Baraer, M., Mark, B. G., McKenzie, J. M., Condom, T., Bury, J., Huh, K. I., Portocarrero, C., Gomez, J., & Rathay, S. (2012). Glacier recession and water resources in Peru's Cordillera Blanca. *Journal of Glaciology*, 58(207), 134–150. http://doi.org/10.3189/2012JoG11J186

Baraer, M., McKenzie, J., Mark, B. G., Gordon, R., Bury, J., Condom, T., Gomez, J., Knox, S., & Fortner, S. K. (2015). Contribution of groundwater to the outflow from ungauged glacierized catchments: A multi-site study in the tropical Cordillera Blanca, Peru. *Hydrological Processes*, 29(11), 2561–2581. http://doi.org/10.1002/hyp.10386

Barnett, T. P., Adam, J. C., & Lettenmaier, D.P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066), 303–309. http://doi.org/10.1038/nature04141

Burns, P., Mark, B., & McKenzie, J. (2011). A multi-parameter hydrochemical characterization of proglacial runoff, Cordillera Blanca, Peru. *The Cryosphere Discussions*, 5, 2483–2521. http://doi.org/10.5194/tcd-5-2483-2011

Bury, J. T., Mark, B. G., McKenzie, J. M., French, A., Baraer, M., Huh, K. I., Luyo, M.A.Z., Lopez, R.J.G. (2011). Glacier recession and human vulnerability in the Yanamarey watershed of the Cordillera Blanca, Peru. *Climatic Change*, 105(1), 179–206. http://doi.org/10.1007/s10584-010-9870-1

Bury, J., Mark, B. G., Carey, M., Young, K. R., Mckenzie, J. M., Baraer, M., French, A., & Polk, M. H. (2013). New geographies of water and climate change in Peru: coupled natural and social transformations in the Santa River watershed. *Annals of the Association of American Geographers*, 103(October 2012), 363–374. http://doi.org/10.1080/00045608.2013.754665

Chavez, D. (2013). Groundwater potential of pampa aquifers in two glacial watersheds, Cordillera Blanca, Peru. *Unpublished M.Sc. Thesis*, McGill University, Montreal, Quebec, Canada.

Clow, D. W., Schrott, L., Webb, R., Campbell, D. H., Torizzo, A., & Dornblaser, M. (2003). Ground Water Occurrence and Contributions to Streamflow in an Alpine Catchment, Colorado Front Range. *Ground Water*, 41(7), 937–950. https://doi.org/10.1111/j.1745-6584.2003.tb02436.x

Glas, R., Lautz, L., McKenzie, J., Mark, B., Baraer, M., Chavez, D., & Maharaj, L. (2018). A review of the current state of knowledge of proglacial hydrogeology in the Cordillera Blanca, Peru. *WIREs Water*, (August), e1299. https://doi.org/10.1002/wat2.1299

Glas, R. L. (2018). Surface water and groundwater in a changing climate: Applications of

hydrogeophysics and trend analysis to understand hydrologic systems. *Doctoral Dissertation*. Syracuse University, Syracuse, New York.

Gordon, R. P., Lautz, L. K., McKenzie, J. M., Mark, B. G., Chavez, D., & Baraer, M. (2015). Sources and pathways of stream generation in tropical proglacial valleys of the Cordillera Blanca, Peru. *Journal of Hydrology*, 522, 628–644. http://doi.org/10.1016/j.jhydrol.2015.01.013

Harbaugh, A.W. (2005). MODFLOW-2005, the U.S. Geological Survey modular ground-water model -- the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16.

Hood, J. L., Roy, J. W., & Hayashi, M. (2006). Importance of groundwater in the water balance of an alpine headwater lake. *Geophysical Research Letters*, 33(13), 1–5. https://doi.org/10.1029/2006GL026611

Kaser, G., Ames, A., & Zamora, M. (1990). Glacier fluctuations and climate in the Cordillera Blanca, Peru. *Annals of Glaciology*, 14, 136–140.

Kaser, G. (1999). A review of the modern fluctuations of tropical glaciers. *Global and Planetary Change*, 22(1–4), 93–103. https://doi.org/10.1016/S0921-8181(99)00028-4

Kaser, G., Juen, I., Georges, C., Gómez, J., & Tamayo, W. (2003). The impact of glaciers on the runoff and the reconstruction of mass balance history from hydrological data in the tropical Cordillera Bianca, Perú. *Journal of Hydrology*, 282(1–4), 130–144. https://doi.org/10.1016/S0022-1694(03)00259-2

Mark, B. G., & Seltzer, G. O. (2003). Tropical glacier meltwater contribution to stream discharge: A case study in the Cordillera Blanca, Peru. *Journal of Glaciology*, 49(165), 271–281. http://doi.org/10.3189/172756503781830746

Mark, B. G., & McKenzie, J. M. (2007). Tracing increasing tropical Andean glacier melt with stable isotopes in water. *Environmental Science and Technology*, 41(20), 6955–6960. https://doi.org/10.1021/es071099d

McClymont, A. F., Roy, J. W., Hayashi, M., Bentley, L. R., Maurer, H., & Langston, G. (2011). Investigating groundwater flow paths within proglacial moraine using multiple geophysical methods. *Journal of Hydrology*, 399(1–2), 57–69. https://doi.org/10.1016/j.jhydrol.2010.12.036

McClymont, A. F., Hayashi, M., Bentley, L. R., Muir, D., & Ernst, E. (2010). Groundwater flow and storage within an alpine meadow-talus complex. *Hydrology and Earth System Sciences*, 14(6), 859–872. https://doi.org/10.5194/hess-14-859-2010

Muir, D. L., Hayashi, M., & McClymont, A. F. (2011). Hydrological storage and transmission characteristics of an alpine talus. *Hydrological Processes*, 25(19), 2954–2966. https://doi.org/10.1002/hyp.8060

Roy, J. W., & Hayashi, M. (2009). Multiple, distinct groundwater flow systems of a single moraine-talus feature in an alpine watershed. *Journal of Hydrology*, 373(1–2), 139–150. https://doi.org/10.1016/j.jhydrol.2009.04.018

Schauwecker, S., Rohrer, M., Acuña, D., Cochachin, A., Dávila, L., Frey, H., Giraldez, C., Gomez, J., Huggel, C., Jacques-Coper, M., Loartes, E., Salzmann, N., & Vuille, M. (2014). Climate trends and glacier retreat in the Cordillera Blanca, Peru, revisited. *Global and Planetary Change*, 119,85–97. https://doi.org/10.1016/j.gloplacha.2014.05.005

Smith, L., & Wheatcraft, S.W. (1993). “Groundwater Flow”. In Handbook of Hydrology Edited by: Maidment, D. R. Chapter 6. McGraw-Hill.

Somers, L. D., Gordon, R. P., McKenzie, J. M., Lautz, L. K., Wigmore, O., Glose, A. M., Glas,

R., Aubry-Wake, C., Mark, B., Baraer, M., & Condom, T. (2016). Quantifying groundwater-surface water interactions in a proglacial valley, Cordillera Blanca, Peru. *Hydrological Processes*, 30(17), 2915–2929. https://doi.org/10.1002/hyp.10912

Suarez, W., Chevallier, P., Pouyaud, B., & Lopez, P. (2008). Modelling the water balance in the glacierized Parón Lake basin (White Cordillera, Peru). *Hydrological Sciences Journal*, 53(1), 266–277. http://doi.org/10.1623/hysj.53.1.266

Vincent, A., Violette, S., & Aðalgeirsdóttir, G. (2019). Groundwater in catchments headed by temperate glaciers: A review. *Earth-Science Reviews*, 188: 59-76. https://doi.org/10.1016/j.earscirev.2018.10.017

Vuille, M., Francou, B., Wagnon, P., Juen, I., Kaser, G., Mark, B. G., & Bradley, R. S. (2008). Climate change and tropical Andean glaciers: Past, present and future. *Earth-Science Reviews*, 89(3–4), 79–96. https://doi.org/10.1016/j.earscirev.2008.04.002

Vuille, M., Kaser, G., & Juen, I. (2008). Glacier mass balance variability in the Cordillera Blanca, Peru and its relationship with climate and the large-scale circulation. *Global and Planetary Change*, 62(1–2), 14–28. https://doi.org/10.1016/j.gloplacha.2007.11.003

Waterloo Hydrogeologic Inc. (2005). Visual MODFLOW v. 4.1. Professional Edition, User's Manual, Ontario, Canada.

Wigmore, O. & Mark, B. (2018). High altitude kite mapping: evaluation of kite aerial photography (KAP) and structure from motion digital elevation models in the Peruvian Andes. *International Journal of Remote Sensing*, 39: 4995-5015. https://doi.org/10.1080/01431161.2017.1387312

Tables

Table 1. Borehole depths and piezometer details for Quilcayhuanca groundwater wells.

Well	TOC Elevation m asl	Well Depth m BGS	Depth to Pressure Transducer m BGS	Water Level m BTOC	Time Water Level Measured**	Screened Unit	Depth to Aquifer m BTOC	K†† m/s
P1	3933.217	3.70	3.27	0.55	13:28	Talus/till	2.439	7.3×10^{-5}
P2*	3930.344	6.17	-	Artesian	-	Talus/till	5.183	-
P3	3932.525	1.51	1.33	0.55	12:40	Talus/till	0.610	2.0×10^{-5}
P4†	3929.409	2.27	2.03	0.82	12:04	Talus/till	1.448	6.5×10^{-6}
P5	3931.952	4.40	3.05	0.58	14:37	Talus/till	4.05	-
P6	3930.486	6.97	5.05	0.25	14:17	Talus/till	6.20	-
P7	3934.430	5.00	4.04	0.82	13:10	Talus/till	4.60	-

m asl – meters above sea level

m BGS – meters below ground

m BTOC – meters below top of casing

TOC – Top of casing

* Well plugged after drilling due to artesian conditions

** As measured on July 22, 2015

†Elevation measurements from Chavez (2013) elevation survey instead of Wigmore & Mark (2018) gps survey

†† Hydraulic conductivity (K) measurements from Chavez (2013)

Table 2. Hydrologic parameters and model inputs used in the optimal model simulation (run T).

Parameter	Value
Horizontal grid size	10.073 m x 10.089 m
Confining Unit Specific yield (S_y)	0.2
Aquifer Unit Specific yield (S_y)	0.3
Confining Unit Effective porosity (n_e)	0.3
Aquifer Unit Effective porosity (n_e)	0.2
Confining Layer Specific Storage (S_s)	$1x10^{-3}$ m^{-1}
Aquifer Unit Specific Storage (S_s)	$1x10^{-4}$ m^{-1}
Confining Layer K_h	$5x10^{-6}$ m/s
Aquifer Layer K_h	$5x10^{-5}$ m/s
Unit K_h/K_v anisotropy	5:1
Drain Conductance	Proportional to K_h (88 m^2/day)
Talus slope recharge (percentage of monthly precipitation volume)	50% in wet season, 70% in dry season, 2 month lag

Table 3. Estimates of spring and groundwater contribution to the Quilcay stream based on chemical mixing: $f_{GWin} = \frac{[Ion]_{downstream} - [Ion]_{upstream}}{[Ion]_{GW} - [Ion]_{upstream}}$. The mean and median ion concentrations in the tributaries and groundwater were compared to the upstream and downstream concentrations.

Ion	**Flux to Stream (% of streamflow)**	
	Mean	Median
K	20.7	21.8
Na	2.7	2.6
Mg	13.7	13.1
SO_4	13.3	13.4
$\delta^{18}O$	15.2	12.3

Table 4. Sensitivity of the MODFLOW simulations to ranges in the input values that reflect the magnitude of uncertainty in the true value of the input.

Parameter	Value 1	Value 2	Ave Head Change V1→V2 (m)	Ave Flux Change V1→V2 (%)
Confining Layer K	$1x10^{-5}$ m/s	$1x10^{-4}$ m/s	-0.86	3.4
Aquifer K	$5x10^{-5}$ m/s	$1x10^{-4}$ m/s	-0.75	3.4
K Anisotropy	5:1	10:1	2.49	-2.5
Confining Layer Thickness	3 m	4 m	0.18	-0.8
Aquifer Thickness	12 m	27 m	-0.25	9.7
Valley Side Recharge	50%	70%	0.32	38
Surface Recharge	10%	50%	0.06	5.1
Drain Conductance	88 m^2/day	0.8 m^2/day	1.47	-0.5
S_y (%)	0.2/0.3	0.1/0.2	-0.02	-0.3
n, n_e (%)	0.4/0.3, 0.3/0.2	0.3/0.2, 0.2/0.1	-0.0008	-0.02
S_s (m^{-1})	.001/.0001	.0001/.00001	-0.002	-0.03

Table 5. Median and maximum travel times of particles released into the model through side talus slope recharge for a subset of explored model configurations.

Run	Max Residence Time (years)	Median Residence Time (days)
A	8.3	772
B	10.5	716
C	2.9	121
D	1.5	75
E	1.2	65
F	3.05	101
G	3.1	104
H	1.2	62
I	1.2	61
J	3.7	332
K	2.9	133
L	0.27	21
M	0.87	35
N	0.87	35
O	0.87	35
P	0.89	35
Q	1.8	83
R	1.0	32
S	3.0	137
T	1.5	85
U	0.45	25
V	2.4	138
W	1.3	62
X	1.5	84
Y	0.9	35
Z	1.2	63
AA	3.9	97

Table S1. Configurations and error metrics of a subset of Modflow simulations, in order of increasing average monthly hydraulic head RMSE.

Run	Layers	Side Flux	Surface Recharge	Confining K (m/s)		Aquifer K (m/s)		Ss (C,A)	Sy (C,A)	n/ne	Flux to Drains + Stream (m^3/day/km)		Head RMSE	Slope of Meas vs Modeled	R^2
				Kv	Kh	Kv	Kh				1-Jul	1-Aug	plus STDEV (m)	plus STDEV	
A	3, 30 m	30% talus, 2 month lag	50%	5*10^-7	1*10^-6	5*10^-6	1*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	877	655	1.13 ± 0.73	0.99 ± 0.05	0.86 ± 0.07
B	3, 30 m	30% talus, 2 month lag	50%	5*10^-7	1*10^-6	1*10^-6	1*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	760	525	1.19 ± 0.68	0.91 ± 0.04	0.89 ± 0.06
C	3, 30 m	30 % wet, 70% dry	50%	5*10^-7	1*10^-6	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	2302	1510	1.21 ± 0.73	0.78 ± 0.06	0.83 ± 0.07
D	3, 30 m	50%, 2 month lag with smear	10%	5*10^-6	1*10^-5	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	3184	1969	1.28 ± 0.90	0.97 ± 0.07	0.87 ± 0.08
E	3, 15 m	30%, 2 month lag	50%	1*10^-6	1*10^-5	1*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	2206	1325	1.28 ± 0.98	0.88 ± 0.07	0.80 ± 0.08
F	4, 30 m	30%, 2 month lag	50%	5*10^-7	1*10^-6	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	2306	1262	1.32 ± 0.78	0.75 ± 0.06	0.82 ± 0.09
G	4, 30 m	30%, 2 month lag	10%	5*10^-7	1*10^-6	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	2266	1235	1.34 ± 0.81	0.74 ± 0.07	0.82 ± 0.09
H	3, 30 m	70%, 2 month lag	50%	1*10^-5	1*10^-4	1*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	5195	2707	1.34 ± 1.02	0.88 ± 0.09	0.82 ± 0.09
I	3, 30 m	70%, 1 month lag	50%	1*10^-5	1*10^-4	1*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	2726	1849	1.35 ± 1.04	0.88 ± 0.10	0.82 ± 0.11
J	3, 30 m	20%, 2 month lag	40%	5*10^-7	1*10^-6	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	1666	1030	1.36 ± 0.97	0.72 ± 0.07	0.76 ± 0.07
K	3, 30 m	30%, 2 month lag	50%	5*10^-7	1*10^-6	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	2339	1226	1.37 ± 0.85	0.79 ± 0.10	0.81 ± 0.09
L	3, 15 m	50%, 2 month lag	50%	1*10^-6	5*10^-6	5*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.3, 0.2/0.2, 0.1	3250	1663	1.38 ± 0.88	0.98 ± 0.13	0.84 ± 0.08
M	3, 15 m	50%, 2 month lag	50%	1*10^-6	5*10^-6	5*10^-5	1*10^-4	0.0001, 0.00001	0.2, 0.3	0.4, 0.3/0.3, 0.2	3251	1658	1.38 ± 0.88	0.98 ± 0.13	0.84 ± 0.08
N	3, 15 m	50%, 2 month lag	50%	1*10^-6	5*10^-6	5*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	3253	1663	1.38 ± 0.88	0.98 ± 0.13	0.84 ± 0.08
O	3, 15 m	50%, 2 month lag	50%	1*10^-6	1*10^-5	1*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	3277	1680	1.38 ± 0.89	0.98 ± 0.12	0.84 ± 0.08
P	3, 15 m	50%, 2 month lag	50%	1*10^-6	5*10^-6	5*10^-5	1*10^-4	0.001, 0.0001	0.1, 0.2	0.4, 0.3/0.3, 0.2	3183	1483	1.40 ± 0.91	0.97 ± 0.13	0.84 ± 0.09
Q	3, 30 m	50%, 2 month lag	50%	1*10^-6	1*10^-5	1*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	4169	2032	1.41 ± 0.92	0.76 ± 0.10	0.80 ± 0.08
R	3, 15 m	70%, 2 month lag	50%	1*10^-6	5*10^-6	5*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	4381	2238	1.43 ± 0.90	1.01 ± 0.09	0.86 ± 0.07
S	3, 50 m	50%, 2 month lag	50%	1*10^-6	5*10^-6	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	4348	1960	1.46 ± 0.83	0.79 ± 0.11	0.83 ± 0.09
T	3, 30 m	50 % wet, 70% dry, 2 month lag	50%	1*10^-6	5*10^-6	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	4443	1977	1.49 ± 0.91	0.99 ± 0.17	0.86 ± 0.07
U	3, 15 m	70%, 2 month lag	50%	1*10^-6	1*10^-5	1*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	4348	2070	1.54 ± 0.97	1.02 ± 0.12	0.85 ± 0.07
V	3, 30 m	30%, 2 month lag	50%	5*10^-7	1*10^-6	5*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	2450	1506	1.57 ± 1.11	0.81 ± 0.08	0.72 ± 0.05
W	3, 30 m	70%, 2 month lag	50%	1*10^-6	1*10^-5	1*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	4827	2262	1.68 ± 0.91	0.81 ± 0.14	0.83 ± 0.09
X	3, 30 m	50%, 2 month lag	50%	5*10^-7	1*10^-6	5*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	3552	1848	1.77 ± 0.98	0.71 ± 0.12	0.78 ± 0.07
Y	3, 15 m	50%, 2 month lag	50%	5*10^-7	1*10^-6	1*10^-5	5*10^-5	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	3034	1328	2.13 ± 1.64	1.22 ± 0.26	0.84 ± 0.05
Z	3, 30 m	70%, 2 month lag	50%	1*10^-7	1*10^-6	5*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	4711	2154	2.21 ± 1.16	0.76 ± 0.15	0.82 ± 0.08
AA	3, 30 m	30%, 2 month lag	50%	1*10^-7	1*10^-6	1*10^-5	1*10^-4	0.001, 0.0001	0.2, 0.3	0.4, 0.3/0.3, 0.2	2277	1337	2.53 ± 1.68	0.40 ± 0.09	0.82 ± 0.10

Figures

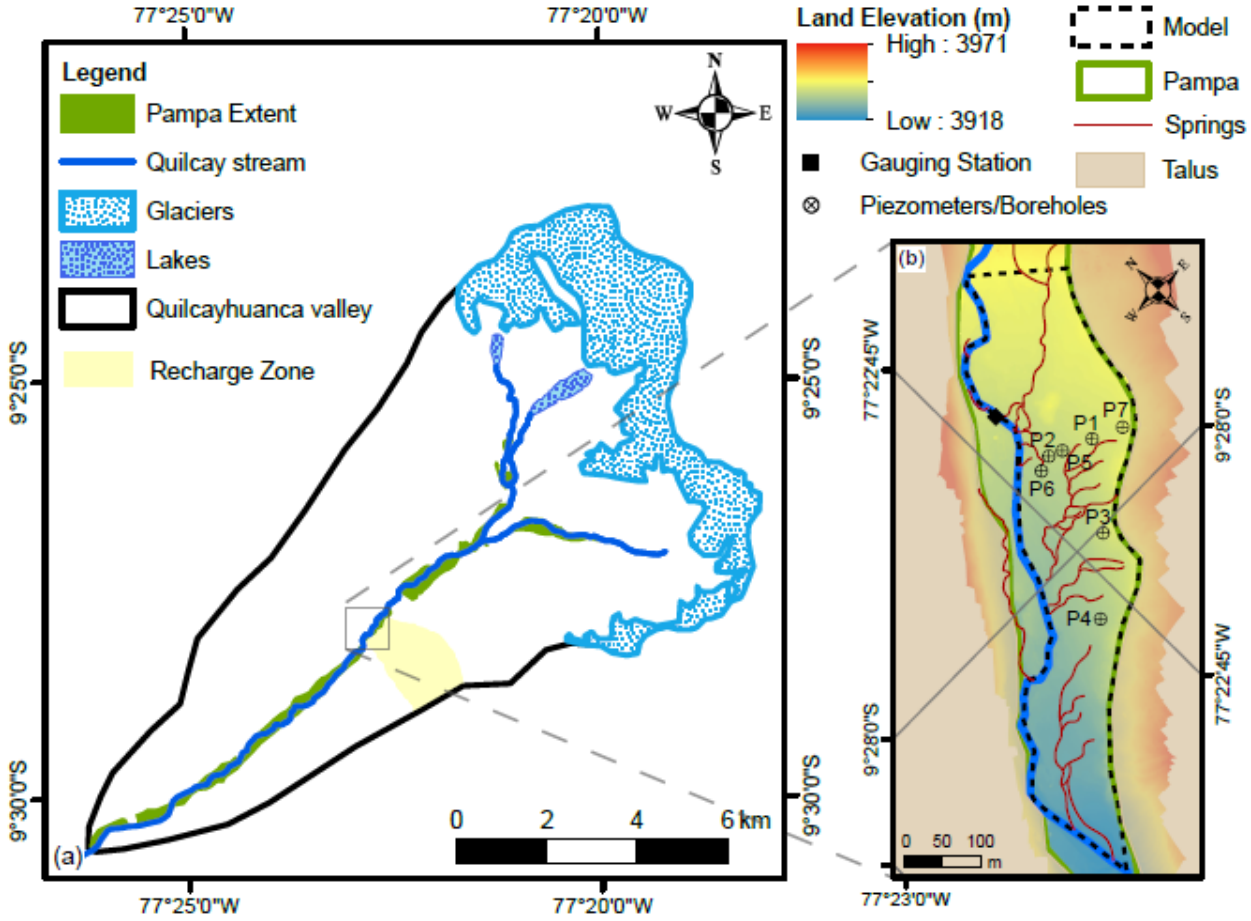

Figure 1. (a) Map of Quilcayhuanca valley depicting the Quilcay stream, pampa extent, and the area from which recharge is sourced. (b) Model domain within the valley showing the locations of the piezometers, talus slope regions, and pampa topography.

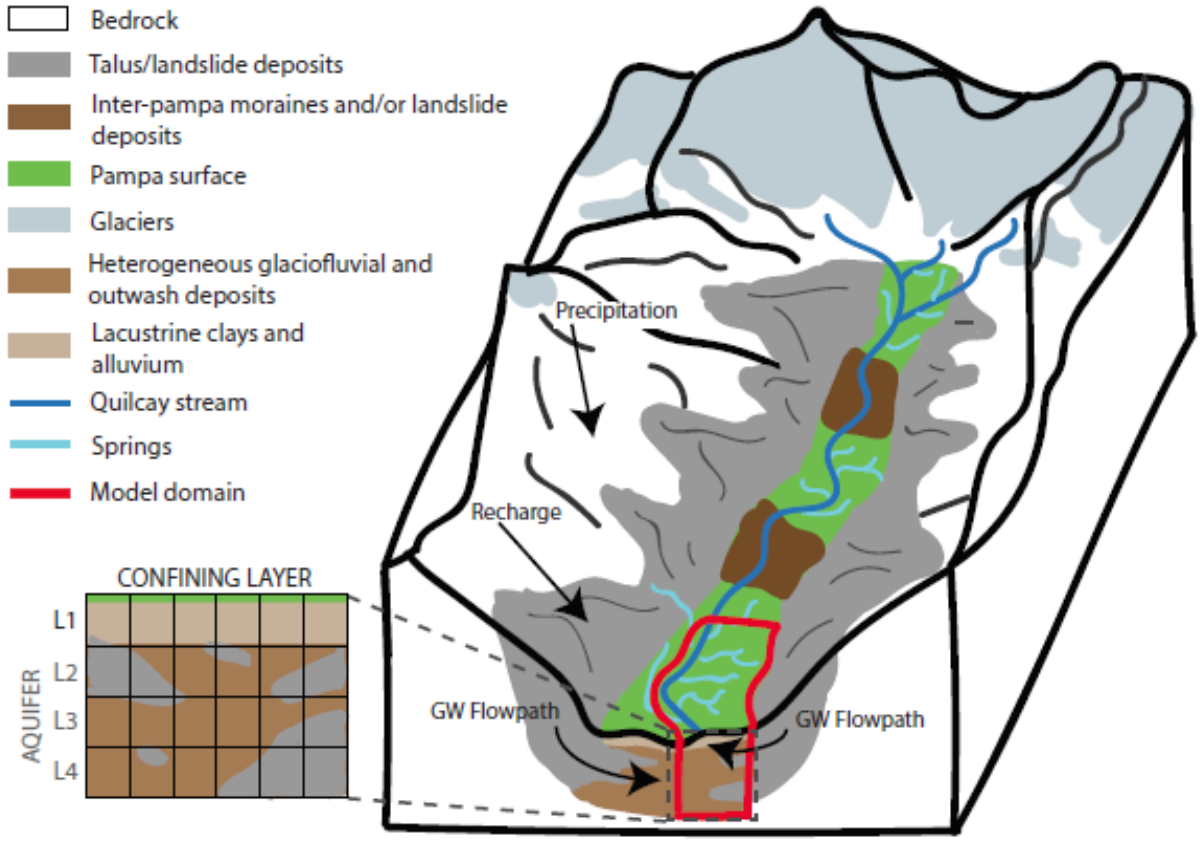

Figure 2. Conceptual model of the pampa aquifer system. Precipitation falls on the bedrock cliffs and infiltrates into the talus slopes/landslide deposits. The groundwater in these side talus deposits recharges the pampa aquifer system. Groundwater flows from the talus deposits into the heterogeneous glaciofluvial and glaciolacustrine deposits beneath the pampa surface. The pampa surface consists of organics, peat, and clays intermixed with sand, serving as a confining layer to the underlying aquifer. Springs occur throughout the pampa, carrying groundwater to the stream.

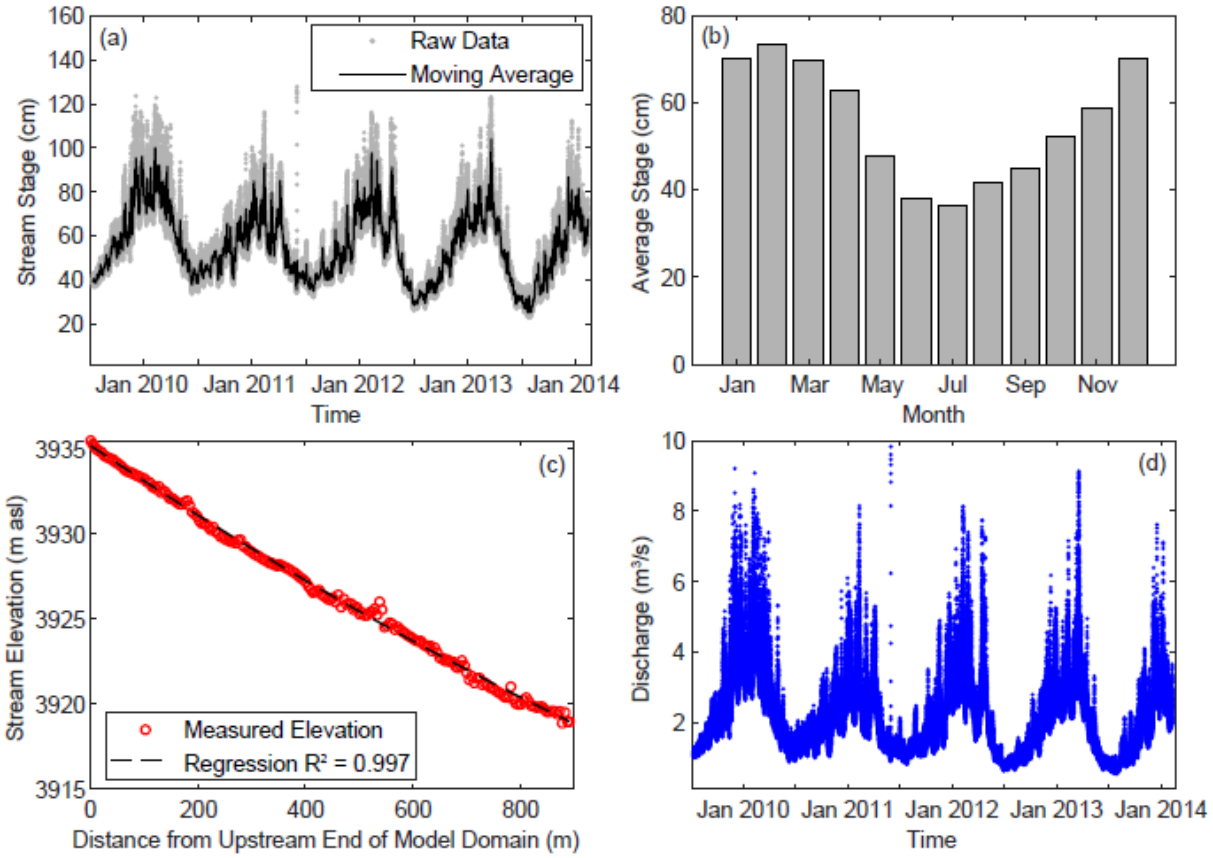

Figure 3. (a) Quilcay stream stage recorded at the gauging station, along with the 100 point (~1 day) moving average. (b) Quilcay stream stage for each month averaged from the ~4.5-year record. (c) Slope of the streambed longitudinally along the Quilcay stream within the model domain. The monthly stream stage and streambed slope are used to create the specified head boundary. (d) Quilcay stream discharge calculated using the rating curve developed at the gauging station.

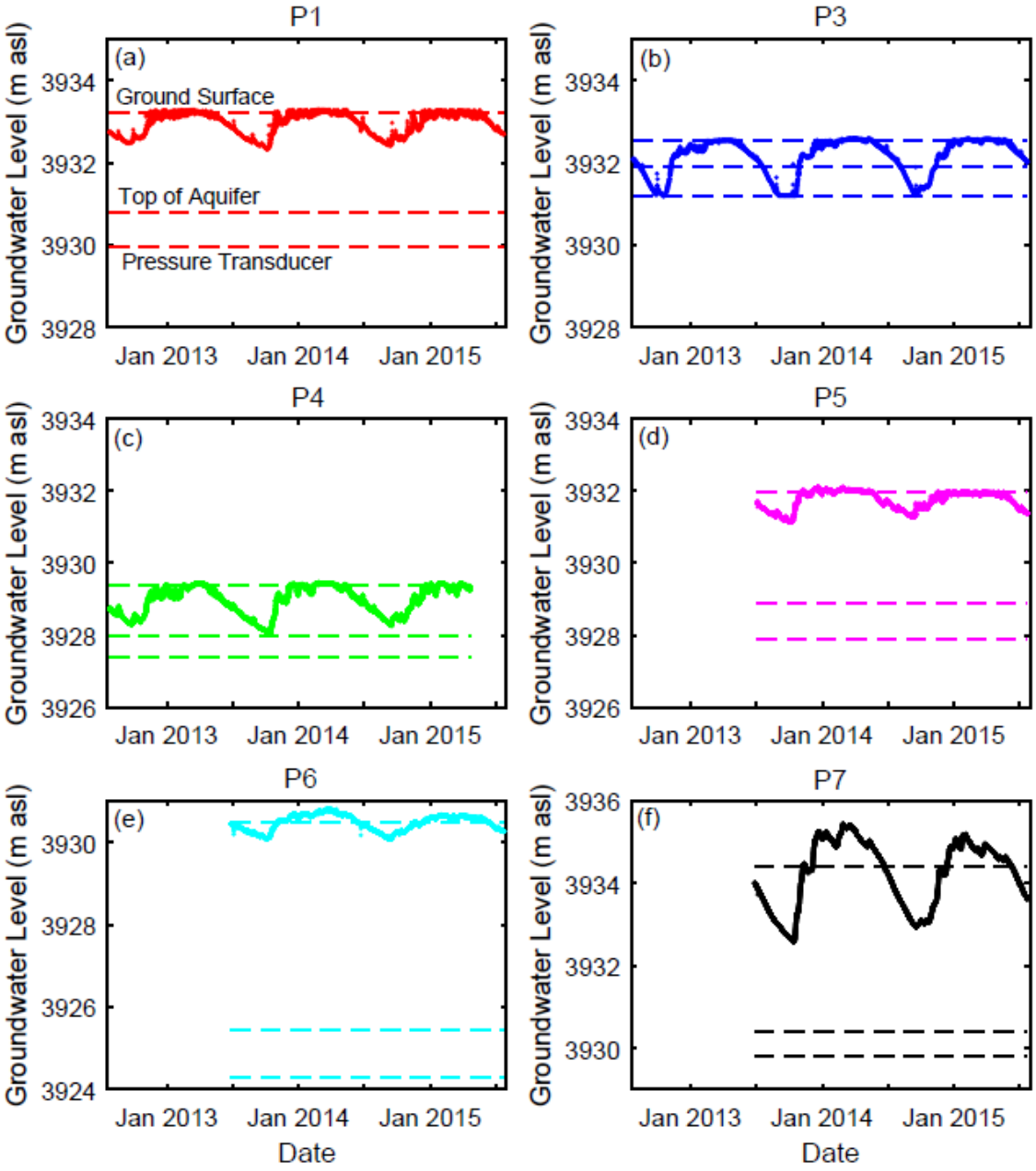

Figure 4. Measured hydraulic heads from the piezometers in the aquifer. The elevations of the ground surface, top of aquifer material, and pressure transducers are marked with dashed lines. The aquifer is over pressurized.

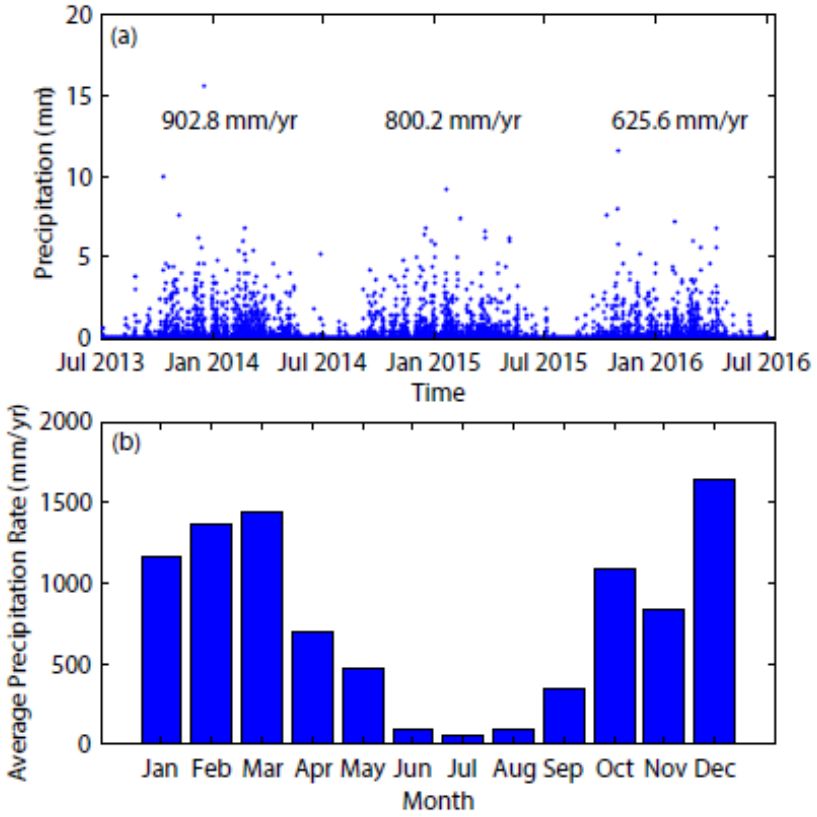

Figure 5. (a) Precipitation measured hourly over a three-year period. (b) Average precipitation rate by month for the three-year period.

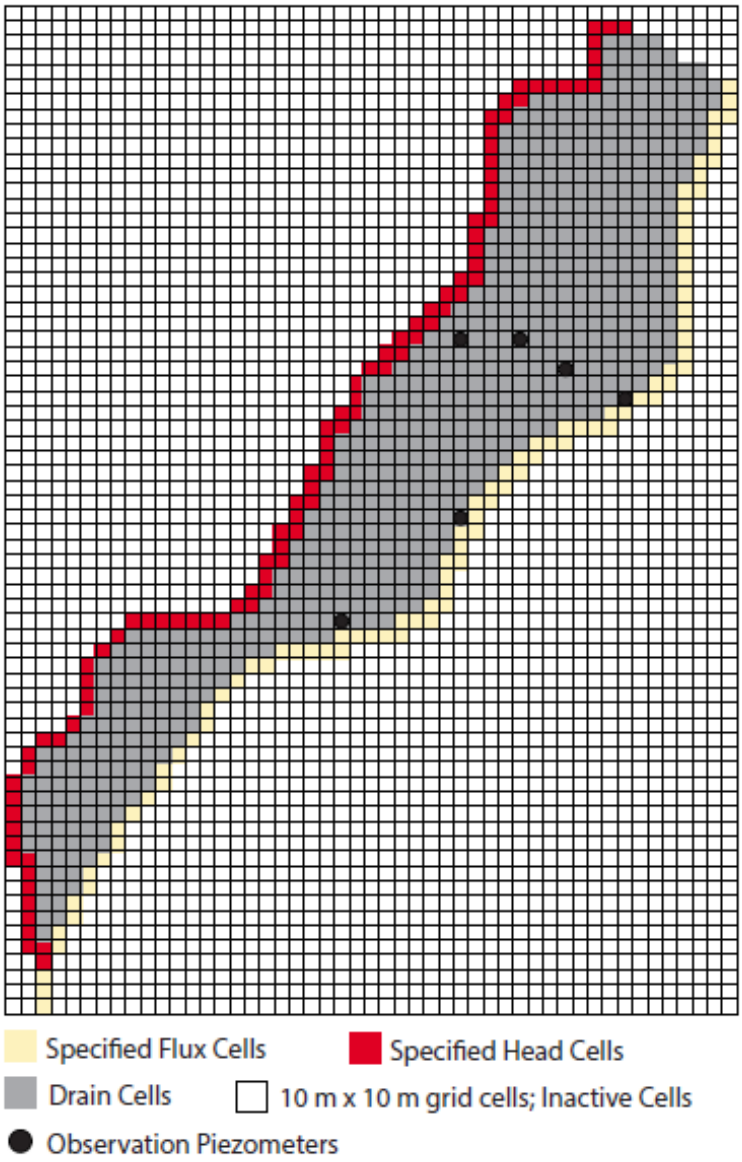

Figure 6. Top cell layer of the model grid and locations of boundary conditions. Specified flux cells simulate the influx of talus slope deposit recharge. The specified head cells simulate the elevation of the stream surface. Drain cells simulate the intermittent occurrence of springs. A no-flow boundary is present at the upper edge of the model reach. Cell layers 2-4 have the same design, but the only boundary condition is the specified flux cells.

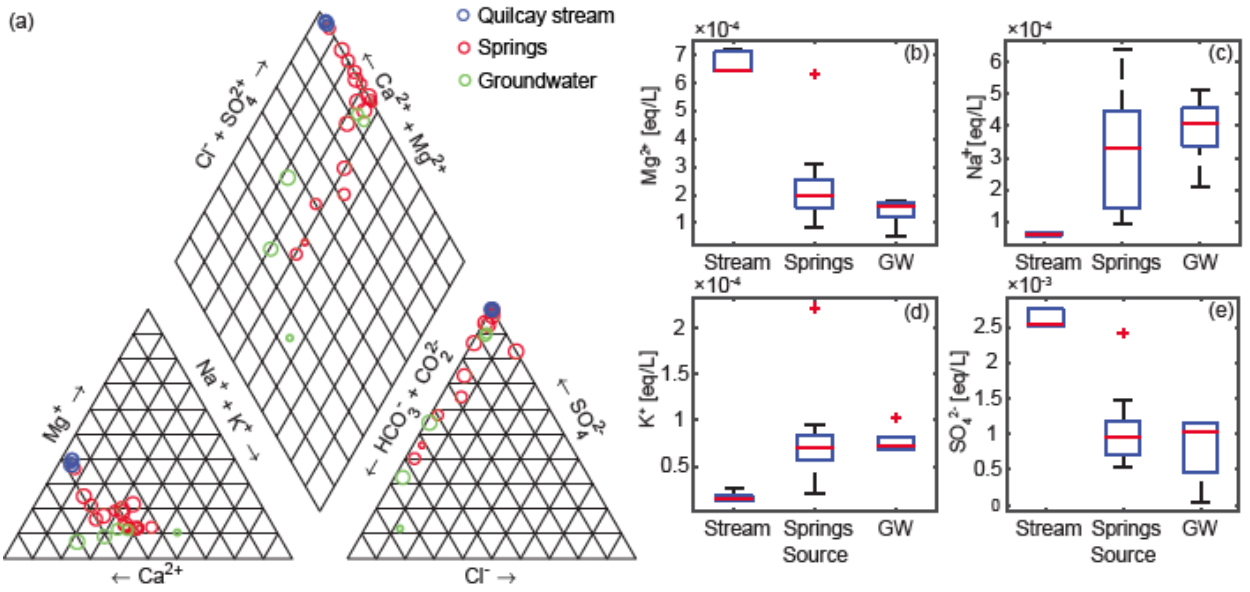

Figure 7. Chemistry of water samples from the Quilcay stream, springs within the pampa, and the piezometers. (a) The symbol size in the piper plot represents the relative ion concentrations in the sample. Sodium (c) and potassium (d) concentrations increased along the study reach, and the samples from the springs and groundwater had higher [K^+] and [Na^+] than the study reach. Meanwhile, magnesium (b) and sulfate (e) concentrations decreased within the model reach, and the spring and groundwater samples contributed water to the stream with lesser amounts of these ions, diluting the concentration in the stream.

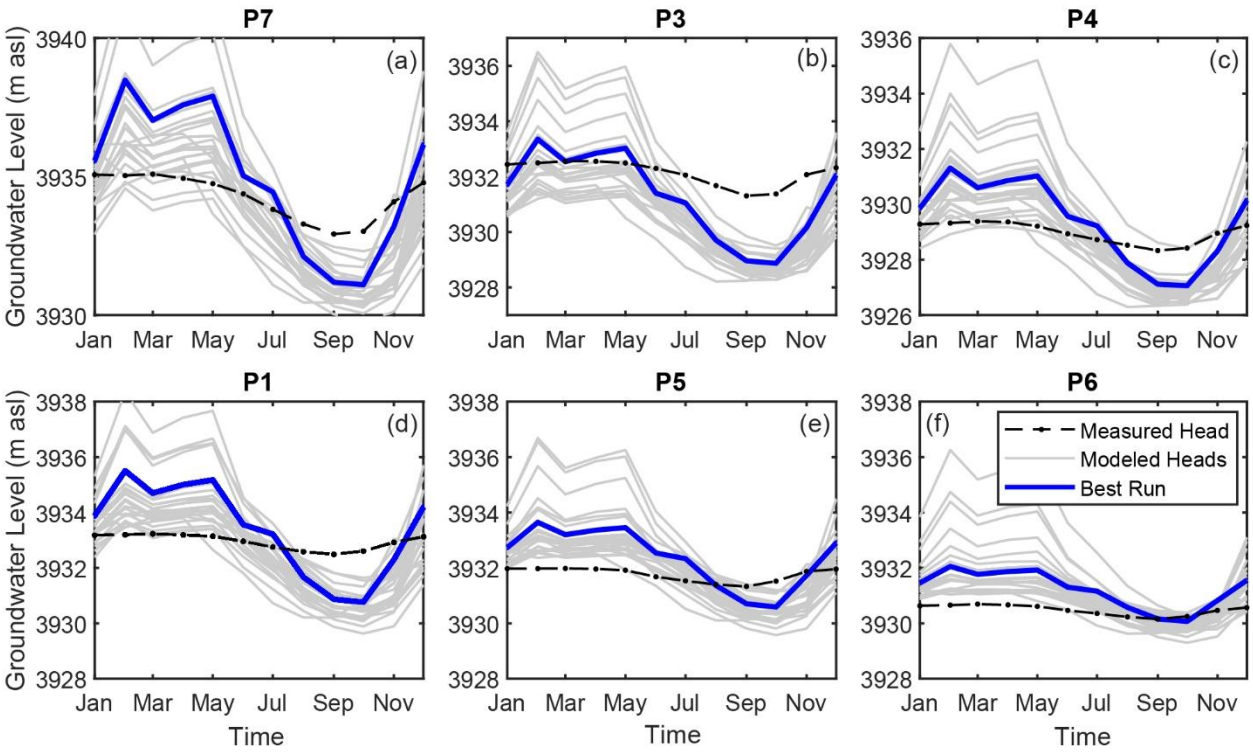

Figure 8. Modeled and measured hydraulic head values through time in the aquifer unit at each piezometer for a subset of MODFLOW simulations with a range of parameter values. The optimal model run (T) is shown in blue.

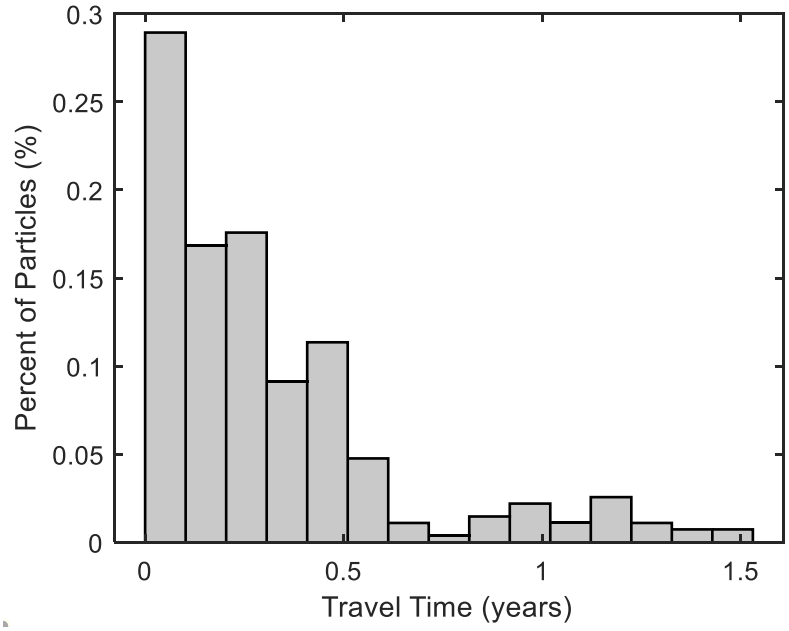

Figure 9. Travel time distribution of 270 particles released into the model through the talus slope recharge boundary. 84% of particle have travel times less than 0.5 years.

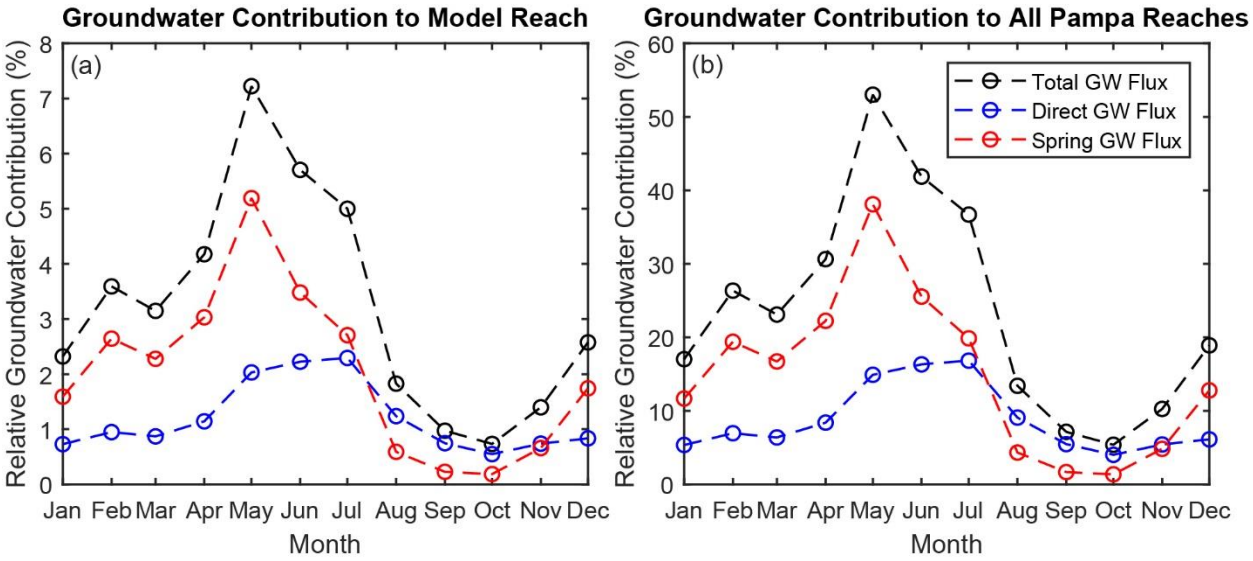

Figure 10. (a) The relative gain in streamflow along the model reach due to groundwater inflow predicted by simulation T (b) The percentage of streamflow in the upper catchment sourced from groundwater that enters the stream within pampa regions of the valley (7.34 km).

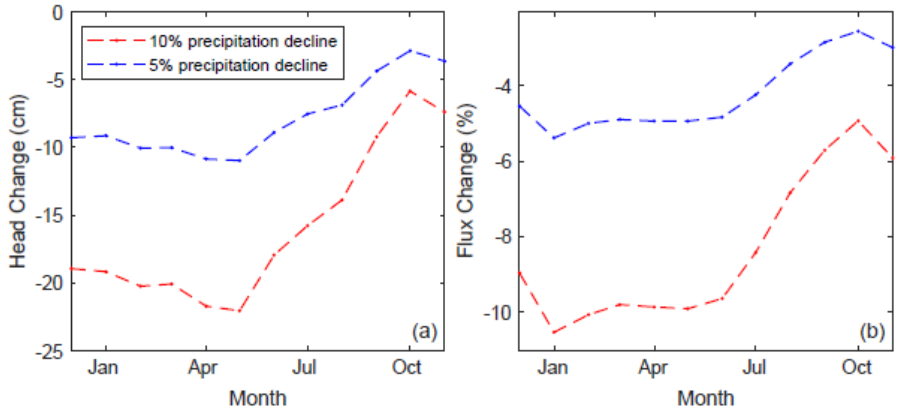

Figure 11. (a) Decline in hydraulic head values at the six piezometers under decreased recharge scenarios. (b) Decrease in total groundwater flux rates due to a one-year decline in recharge.

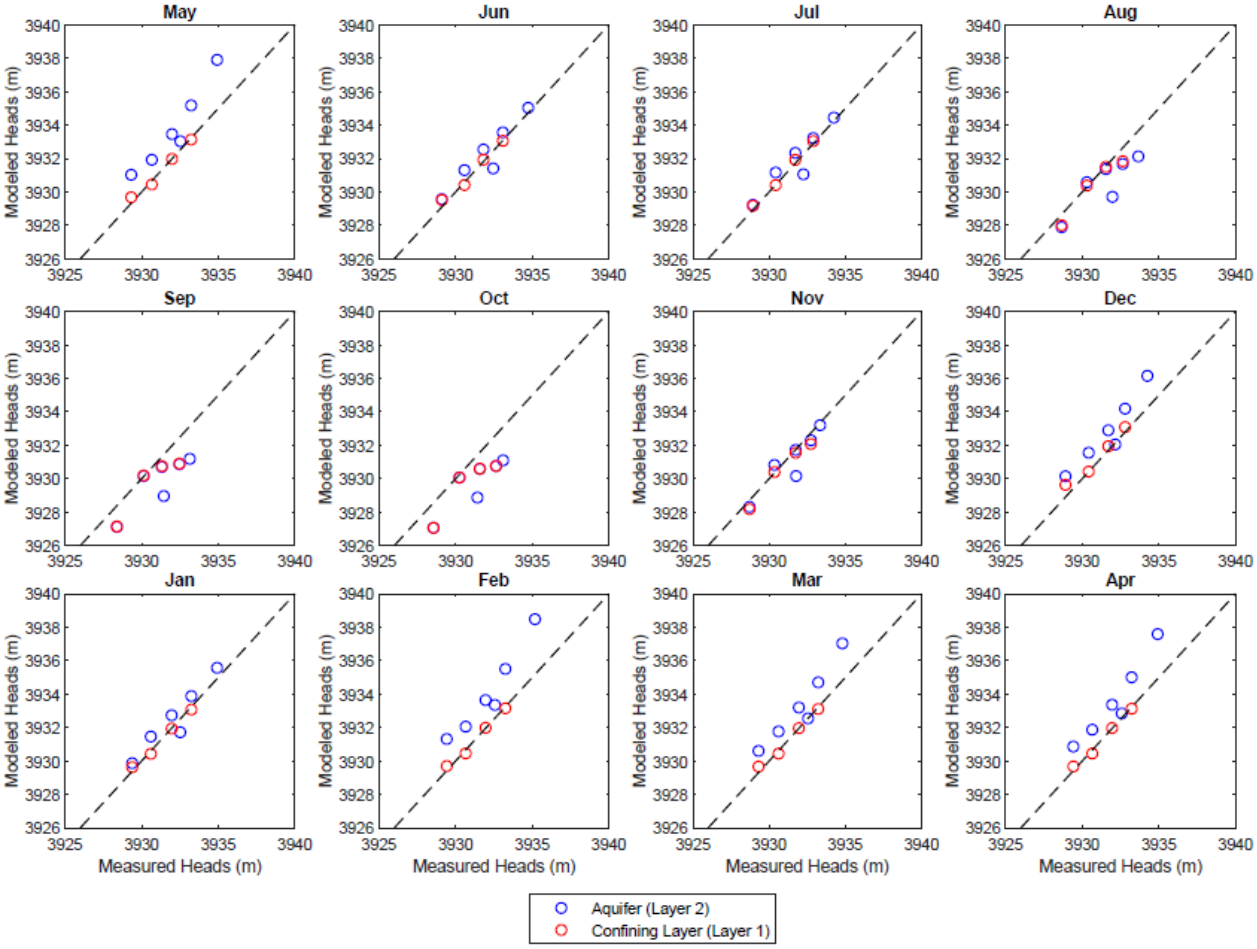

Figure S1. Modeled verses measured hydraulic heads for each month of run T. Heads were not measured in the confining layer and so the heads measured in the aquifer unit were used to compare the modeled heads from the two layers. The black dashed line is the 1:1 line.

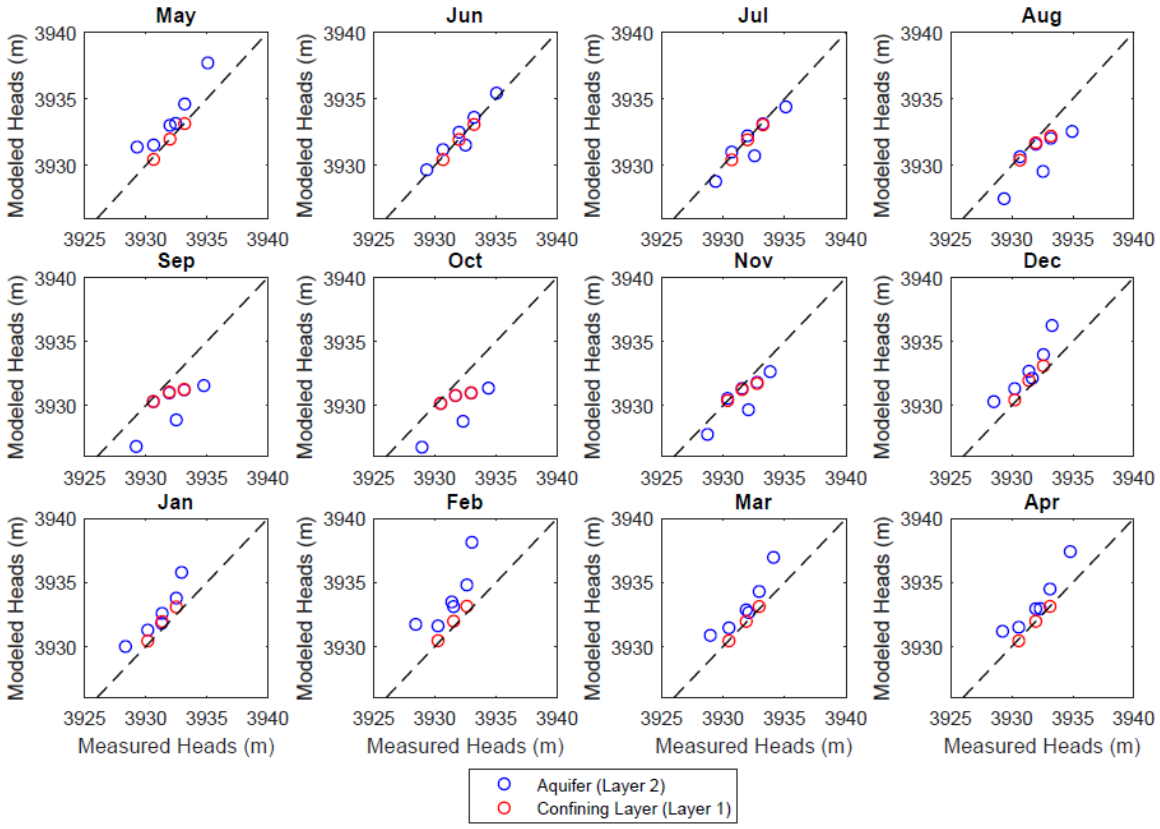

Figure S2. Modeled verses measured hydraulic heads for each month of run R. Heads were not measured in the confining layer and so the heads measured in the aquifer unit were used to compare the modeled heads from the two layers. The black dashed line is the 1:1 line.

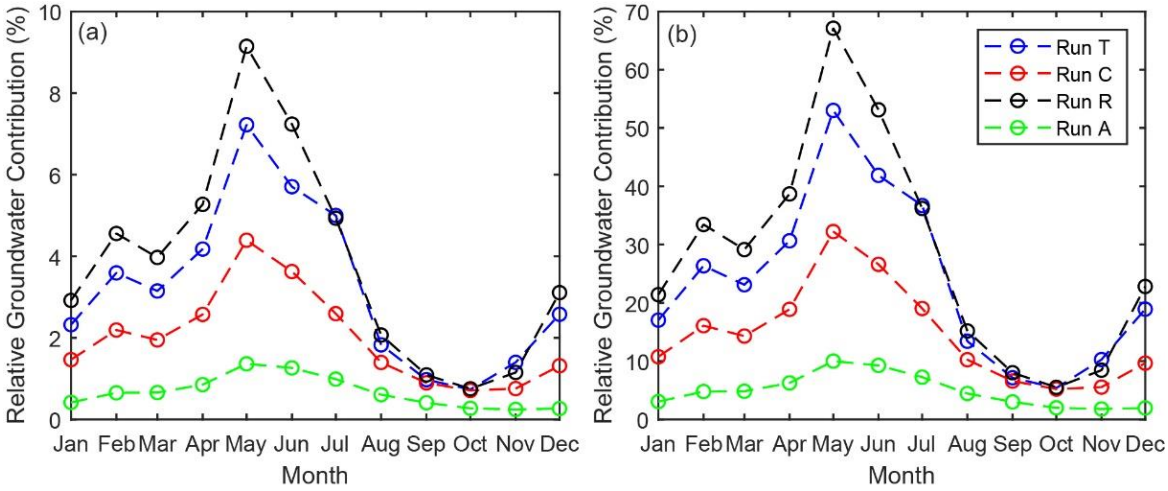

Figure S3. (a) The percent gain in streamflow along the model reach due to total groundwater inflow predicted by a suite of simulations. (b) The percentage of streamflow in the upper catchment sourced from groundwater entering within pampa regions (7.34 km).

Vita

EDUCATION

Ph.D. Earth Sciences *May 2019*
Syracuse University, Syracuse, NY
Dissertation: Surface Water – Groundwater Interactions in a Proglacial Alpine Catchment: Applications of Heat Tracing, Modeling, and Remote Sensing Methods

B.A., Geology & Statistics (minor), Magna cum laude with Honors in Geology *May 2015*
Mount Holyoke College, South Hadley, MA
Honors Thesis: Calcite-Graphite Isotope Thermometry of Marble in the Bancroft Shear Zone

PUBLICATIONS

Baker, E.A., L.K. Lautz, J.M. McKenzie, C. Aubry-Wake. 2019. Improving the accuracy of time-lapse thermal infrared imaging for hydrologic applications, *Journal of Hydrology*, 571, 60-70. https://doi.org/10.1016/j.jhydrol.2019.01.053

Caldwell, S., C. Kelleher, E. **Baker**, L. K. Lautz. 2019. Stream temperature dynamics from above: using thermal infrared imagery to observe and model stream temperature, *Science of the Total Environment*, 661, 364-374. https://doi.org/10.1016/j.scitotenv.2018.12.457

Baker, E.A., L.K. Lautz, C. Kelleher, J.M. McKenzie. 2018. The importance of incorporating diurnally fluctuating stream discharge in stream temperature energy balance models. *Hydrological Processes*, 32, 2901-2914. https://doi.org/10.1002/hyp.13226

Glose, A.M., L.K. Lautz, E.A. **Baker**. 2017. Stream heat budget modeling with HFLUX: model development, verification, and applications across contrasting sites and seasons. *Environmental Modeling & Software*, 92, 213-228. https://doi.org/10.1016/j.envsoft.2017.02.021

RESEARCH EXPERIENCE

PhD Dissertation Research, Syracuse University *June 2015 – May 2019*
Advisor: Dr. Laura Lautz, Department of Earth Sciences
Groundwater Flow Modeling

- Assisted with the maintenance of groundwater level loggers in piezometers installed in a remote alpine aquifer system and downloaded and curated annual data
- Created a groundwater flow model using Visual MODFLOW Flex and used MODPATH to estimate groundwater residence time in an Andean proglacial aquifer system

Water Chemistry Sampling & Analysis

- Planned, organized, and executed collection of 68 water samples collected over 3 days from streams, tributaries, wells, and lakes in a valley in the Peruvian Andes
- Measured pH, temperature, and specific conductivity of stream using a YSI meter
- Assisted with trace metal water sampling of a stream in the Peruvian Andes

- Measured the major cations and anions in the 68 water samples using ion chromatography and ICP-OES, and measured the water isotopes using a Picarro

Stream Temperature Energy Balance Modeling

- Reprogrammed the HFLUX stream temperature model to enable stream discharge to change through time
- Created a spatial and temporal energy balance model of a stream reach in the Quilcayhuanca Valley located in the Cordillera Blanca, Peru to assess groundwater inputs
- Performed uncertainty analyses (Monte Carlo parameter sampling) to evaluate if incorporation of diurnal discharge improves modeled stream temperatures and estimated groundwater inflows

Ground-based Thermal Infrared Methods Analysis

- Recorded and analyzed >1000 time-lapse thermal infrared (TIR) images of stream temperatures in order to assess discrete groundwater inflows along the study reach
- Assessed the current methods used to correct TIR imagery and explored the influence of reflected radiation on TIR temperature data

Undergraduate Thesis Research, Mount Holyoke College *Aug. 2014 – May 2015*
Advisor: Dr. Steven Dunn, Department of Geology
Objective: Determine whether extensional shearing in the Bancroft Shear Zone re-set the calcite-graphite thermometer to the temperature at the time of shearing to constrain conditions that re-set this thermometer.

- Collected mylonitic marble samples from the Bancroft Shear Zone
- Prepared samples for isotopic analysis by powdering the samples, separating the graphite from the calcite, and combusting the graphite to form carbon dioxide
- Extracted and purified CO_2 using a vacuum line in the stable isotope laboratory
- Analyzed isotopic data both graphically and statistically in R to assess whether the calcite-graphite thermometer was reset
- Wrote and defended research as a senior honors thesis

Marine Biogeochemistry REU Intern, University of South Carolina *May – Aug. 2014*
Research Advisor: Dr. Seth John, Department of Earth and Ocean Sciences
Objective: Determine the effects of temperature and pH on the isotopic fractionation as Fe(III) is chemically and photochemically reduced to Fe(II).

- Conducted laboratory experiments to reduce Fe(III) to Fe(II) both chemically and photochemically at four temperature and five pH conditions
- Measured Fe(II) concentration using a spectrophotometer and separated Fe(II) from Fe(III) using a peristaltic pump and carbon-18 column
- Prepared Fe(II) samples for isotopic analysis on the mass spectrometer through anion exchange chromatography in the trace metal clean laboratory

Research Intern, Cary Institute of Ecosystem Studies, Millbrook, NY *May – Aug. 2013*
Research Advisor: Dr. Shannon LaDeau, Disease Ecologist

- Conducted biweekly sampling of 10 ephemeral pools to collect mosquito larvae

- Determined species of 800 mosquito larvae using a dichotomous key
- Sampled the adult mosquito population around ephemeral pools using carbon dioxide light traps
- Counted 4000+ adult mosquitos and identified the genus of a subsample of 300 adult mosquitos
- Collected leaf litter samples within the ephemeral pools when they were dry and soaked these samples to determine how many viable mosquito eggs were within the dry ephemeral pools

Research Advisor: Victoria Kelly, Environmental Monitoring Program Manager

- Compiled dataset of the inorganic contents of 900 private wells from an online database
- Input data into GIS so effects of road salt usage on groundwater could be analyzed
- Filtered water samples from a stream during a flood event using vacuum filtration and measured the pH, conductivity, and turbidity of each water sample
- Determined the mass of organic and inorganic materials within each water sample using a muffle furnace to remove the organic materials

TECHNICAL SKILLS

Software: MATLAB, R, Visual MODFLOW, ArcGIS, AquaChem, AQTESOLV, Adobe Illustrator, Excel

Laboratory: ICS 2000 Ion Chromatograph (IC), Picarro L2130-I Water Isotope Analyzer

Field Equipment: Marsh-McBirney Flow Meter, Total Station, HOBO Water Level Data Loggers, iButton Temperature Loggers, Jenoptik HD Thermal Infrared Camera, Vantage Pro2 Meteorological Station, YSI pH/conductivity meter

CONSULTING EXPERIENCE

Environmental Consulting Intern, Geosyntec Consultants, Seattle, WA *Summer 2018*

- Worked as an environmental consulting intern for ~40 hours per week for 8 weeks
- Contaminated site data evaluation
- Technical report preparation
- Fieldwork support (groundwater sampling, pump tests, slug tests, site data collection)

CERTIFICATIONS

OSHA 40-Hour HAZWOPER *April 2018*

TEACHING EXPERIENCE

Teaching Assistant, Syracuse University, Syracuse, NY

EAR 117: Oceanography *Spring 2019*

- Taught four recitation sections and graded weekly assignments
- Served as TA coordinator

EAR 205: Water and Our Environment *Fall 2018*
- Taught four recitation sections and graded weekly assignments
- Gave two lectures to the 100 student class (~1.5 hours each)

EAR 105: Earth Science 105 *Fall 2015*
- Taught four recitation sections and graded weekly assignments

Undergraduate Teaching Assistant, Mount Holyoke College, South Hadley, MA
Geology 322: Igneous & Metamorphic Petrology *Spring 2015*
Geology 201: Rocks & Minerals *Fall 2014*
Statistics 240: Elementary Data Analysis and Experimental Design *Fall 2013 – Fall 2014*
Mathematics 101: Calculus I *Spring 2013*

FELLOWSHIPS

Syracuse University Water Fellowship *Aug. 2017 – May 2018*
Energy Model Program on Water-Energy Research, NSF NRT Program *Sept. 2016 – Aug. 2017*

GRANTS

EMPOWER Seed Grant in support of summer internship *May 2018*
Northeast GSA Travel Grant Recipient *September 2017*
EMPOWER Seed Grant to attend Thermal Infrared Training Course *May 2017*
Northeast GSA Travel Grant Recipient *September 2016*
Northeast GSA Travel Grant Recipient *February 2016*
CNYAPG Grant for Student Research Recipient *January 2016*

AWARDS AND HONORS

Outstanding Student Presentation Award, AGU Annual Conference *Dec 2018*
Chairman's Award, outstanding service to the department and professional promise *May 2017*

PRESENTATIONS

Poster Presentations:
Baker, E.A., L.K. Lautz, and J.M. McKenzie. Improving the accuracy of stream temperatures acquired through ground-based time-lapse thermal infrared imagery. Proceedings of the America Geophysical Union Annual Meeting, December 10-14, 2018: Washington, D.C. H11H-1561.

Baker, E.A. L.K. Lautz, J.M. McKenzie, and B.G. Mark. Methods for correcting ground-based time-lapse infrared imagery. Geological Society of America Abstracts with Programs. Vol. 49, No. 6, doi: 10.1130/abs/2017AM-305398, October 22-25, 2017: Seattle, Washington.

Baker, E.A. L.K. Lautz, C. Kelleher, and J.M. McKenzie. The importance of diurnal fluctuations in stream discharge for determining groundwater inflow. Gordon Research Conference on Catchment Science: Interactions of Hydrology, Biology & Geochemistry, June 25-30, 2017: Lewiston, Maine.

Baker, E.A., L.K. Lautz, J.M. McKenzie, A. Glose, and C. Kelleher. The effect of channel geometry and diurnal discharge fluctuations on modeled stream temperatures. Proceedings of the American Geophysical Union Annual Meeting, December 12-16, 2016: San Francisco, California. H33B-1543.

Baker, EA, LK Lautz, C Aubry-Wake, JM McKenzie, RL Glas, BG Mark. Infrared Imaging and Modeling of Proglacial Stream Temperature in the Cordillera Blanca, Peru. Proceedings of the American Geophysical Union Annual Meeting, December 14-18, 2015: San Francisco, California. H23H-1670.

Oral Presentations:
Baker, E.A., L.K. Lautz, J.M. McKenzie, and A. Glose. How do amplitude and phase shift of diurnal discharge fluctuations affect stream temperature models? Geological Society of America Abstracts with Programs. Vol. 48, No. 7. doi: 10.1130/abs/2016AM-285548, September 25-28, 2016: Denver, Colorado.

Baker, E.A., L.K. Lautz, J.M. McKenzie, C. Aubry-Wake, O. Wigmore, and B.G. Mark. Infrared imaging of proglacial stream temperature in the Cordillera Blanca, Peru. Central New York Association of Professional Geologists Monthly Meeting, September 15, 2016: Syracuse, New York.

Baker, Emily, Laura Lautz, Jeffrey McKenzie, Caroline Aubry-Wake, Lauren Somers, Oliver Wigmore, AnneMarie Glose, Robin Glas, Bryan Mark. Infrared imaging and modeling of proglacial stream temperature in the Cordillera Blanca, Peru. Proceedings of the Northeastern Section of the Geological Society of America, March 21-23, 2016: Albany, New York.

AFFILIATED PRESENTATIONS

Baker, E.A., L.K. Lautz, J.M. McKenzie, C. Kelleher. Illuminating the unseen: resolving how reflection impacts stream temperature observations from time-lapse, ground-based IR cameras. Canadian Geophysical Union Annual Meeting, June 10-14, 2018: Niagara Falls, New York.

McKenzie, J.M., R.L. Glas, L.K. Lautz, B.G. Mark, O. Wigmore, M. Baraer, E.A. **Baker**. Hydrologic transformation of the glacierized watersheds in Peruvian Andes: From glaciers to groundwater. Proceedings of the American Geophysical Union Annual Meeting, December 12-16, 2016: San Francisco, California. H13L-1588.

Glas, R.L., L.K. Lautz, J.M. McKenzie, E.A. **Baker**, L.D. Somers, C. Aubry-Wake, O. Wigmore, B.G. Mark, R. Moucha. Integrating multiple geophysical methods to quantify alpine groundwater-surface water interactions: Cordillera Blanca, Peru. Proceedings of the American Geophysical Union Annual Meeting, Dec. 12-16, 2016: San Francisco, California. NS43C-1937.

Glas, R.L., LK. Lautz, J.M. McKenzie, C. Aubry-Wake, E.A. **Baker**, L. Somers, B.G. Mark and O. Wigmore. Using multiple, integrated hydrogeophysical methods to understand groundwater/surface water interactions in glaciated, tropical catchments of the Cordillera Blanca,

Peru. Geological Society of America Abstracts with Programs. Vol. 48, No. 7 doi: 10.1130/abs/2016AM-285874, September 25-28, 2016: Denver, Colorado.

Glas, R.L., L.K. Lautz, J.M. McKenzie, C. Aubry-Wake, E.A. **Baker**, L. Somers, B.G. Mark, O. Wigmore. Characterization of aquifer structure using seismic refraction tomography in the Cordillera Blanca, Peru. Foro Internacional de Glaciares y Ecosystemas de Montana. August 10-13, 2016: Huaraz, Peru.

McKenzie, J.M., C. Aubry-Wake, E.A. **Baker**, L.K. Lautz, O. Wigmore, M. Baraer, B.G. Mark. Hot and Hotter: Temperature as an indicator of environmental change and a tracer of hydrologic processes. Proceedings of the Canadian Geophysical Union, May 29-June 2, 2016: Fredericton, New Brunswick.

Glas, R.L., L.K. Lautz, J.M. McKenzie, E.A. **Baker**, C. Aubry-Wake, L. Somers, O. Wigmore. Constraining subsurface structure and composition using seismic refraction surveys of proglacial valleys in the Cordillera Blanca, Peru. Proceedings of the Northeastern Section of the Geological Society of America, March 21-23, 2016: Albany, New York.

Glas, R.L., L.K. Lautz, J.M. McKenzie, E.A. **Baker**, C. Aubry-Wake, L. Somers. Constraining Subsurface Structure and Composition Using Seismic Refraction Surveys of Proglacial Valleys in the Cordillera Blanca, Peru. Proceedings of the American Geophysical Union Annual Meeting, December 14-18, 2015: San Francisco, California.

WORKSHOPS & SHORT COURSES

Water – Energy Field Course: International Field Experience, Lake Kivu, Rwanda *June 2018*
Water – Energy Field Course: Domestic Field Experience, Northeastern U.S. *August 2017*
Level I Infrared Thermography Training Course, Infrared Training Center *May 2017*
Applied Geochemical Methods for Mountain Hydrology, McGill University *January 2017*
Practical Techniques for Using Temperature as a Tracer in Hydrological Research *GSA 2016*
AAAS Workshop *Spring 2016*
National Research Traineeship (NRT) Conference *May 2016*

SERVICE & LEADERSHIP

Faculty Representative, Geology Graduate Organization (GeoGo) *Spring 2018*

Syracuse University Geology Club
President *2016-2017*
Vice President *2015-2016*

MEMBERSHIPS

American Geophysical Union (AGU), Member *2014-Present*
Geological Society of America (GSA), Member *2016-Present*
WiSE-FPP (Women in Science and Engineering Future Professionals Program) *2016-2018*